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 **41282694.9km** in 1990. As of 2016, the most recent year for which data was available, that number had fallen to **39958245.9km**, a loss of **1324449km**, or **3,20%**.

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 **1279999.9891km**).

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%**
Sub-Saharan Africa	**30,67%**	**28,79%**
World	**32,42%**	**31,38%**

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 **527229.06km**. 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 **79200km**, 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 & Caribbean**	**541510.00km**
Indonesia	**East Asia & Pacific**	**282193.98km**
Myanmar	**East Asia & Pacific**	**107234.00km**
Nigeria	**Sub-Saharan Africa**	**106506.00km**
Tanzania	**Sub-Saharan Africa**	**102320.00km**

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
1	**85**
2	**72**
3	**38**
4	**9**

The largest number of countries in 2016 were found in the **FIRST** quartile.

There were **85** 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?*

With the analysis above we can conclude that the forest area of the world as a whole decreased a lot between 1990 and 2016. The two major regions responsible for this decrease were Latin America & Caribbean and Sub-Saharan Africa.
On other hand, we have a few countries whose forest area increased between 1990 and 2016. The major ones are China, the United States, and Iceland.
In the end, more countries decreased their forest area rather than increased between 1990 and 2016.

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

There is a big need to focus on countries in regions such as Latin America & Caribbean and Sub-Saharan Africa. These regions hold the most forest area in the world.
The countries that had the most decrease in forest area are in the region of Sub-Saharan Africa. Which are Togo, Nigeria, Mauritania, and Uganda.
Countries located in the region of Latin America & Caribbean also need urgent attention. Those are Brazil and Honduras.

6. APPENDIX (SQL QUERIES)

A. PART 1 - Global Situation

Create a View called “forestation” by joining all three tables - forest_area, land_area and regions in the workspace.

```
DROP VIEW IF EXISTS forestation;

CREATE VIEW forestation AS
 SELECT country_code, country_name, r.region, year, round(f.forest_area_sqkm::numeric, 2) as forest_area_sq_km,
round(l.total_area_sq_mi::numeric, 2) * 2.59 as total_area_sq_km,
 f.forest_area_sqkm/(l.total_area_sq_mi * 2.59) * 100 as forest_percentage
 , r.income_group
 FROM forest_area f
        JOIN land_area l
        USING (country_code, country_name, year)
        JOIN regions r
        USING (country_code, country_name);
```

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 region, year, forest_area_sq_km
FROM forestation
WHERE region = 'World'
ORDER BY year
LIMIT 1;
```

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 region, year, forest_area_sq_km
FROM forestation
WHERE region = 'World'
ORDER BY year DESC
LIMIT 1;
```

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

```
WITH t1 as (
        SELECT forest_area_sq_km as forest16
FROM forestation
WHERE region = 'World' and year = 2016
),

t2 as (
        SELECT forest_area_sq_km as forest90
FROM forestation
WHERE region = 'World' and year = 1990
)
```

```
SELECT t1.forest16, t2.forest90, (t2.forest90 - t1.forest16) as result
FROM t1, t2
```

D - What was the percent change in the forest area of the world between 1990 and 2016?

```
WITH t1 as (
SELECT forest_area_sq_km as forest16
FROM forestation
WHERE region = 'World' and year = 2016
),

t2 as (
SELECT forest_area_sq_km as forest90
FROM forestation
WHERE region = 'World' and year = 1990
)

SELECT (t2.forest90 - t1.forest16) / t2.forest90 * 100 as result
FROM t1, t2
```

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 year, country_name, total_area_sq_km
FROM forestation
WHERE total_area_sq_km <= 1324449 and year = 2016
ORDER BY 3 DESC
LIMIT 1;
```

B. PART 2 - Regional Outlook

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)

```
DROP VIEW IF EXISTS regions2;

CREATE VIEW regions2
AS
SELECT r.region,
      l.year,
      SUM(f.forest_area_sqkm) total_forest_area_sq_km,
      SUM(l.total_area_sq_mi*2.59) AS total_area_sq_km,
       (SUM(f.forest_area_sqkm)/SUM(l.total_area_sq_mi*2.59))*100 AS forest_percentage_region
FROM forest_area f
            JOIN land_area l
                  ON f.country_code = l.country_code AND f.year = l.year
       JOIN regions r
                  ON l.country_code = r.country_code
GROUP BY 1,2
```

```
ORDER BY 1,2;
```

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?

```
WITH t0 as (
        SELECT round(forest_percentage_region::numeric,2) as world_forest_percentage_2016
        FROM regions2
        WHERE region = 'World' and year = 2016
),

t1 as (
        SELECT year, region, MAX(forest_percentage_region) as max_2016
   FROM regions2
   WHERE year = 2016
   GROUP BY 1, 2
        ORDER BY 3 DESC
        LIMIT 1
),

t2 as (
        SELECT year, region, MIN(forest_percentage_region) as min_2016
   FROM regions2
   WHERE year = 2016
   GROUP BY 1, 2
        ORDER BY 3
        LIMIT 1
)

SELECT t0.world_forest_percentage_2016, t1.region as max_region, round(t1.max_2016::numeric,2) as
forest_percentage_max, t2.region as min_region, round(t2.min_2016::numeric,2) as
forest_percentage_min
FROM t0, t1, t2
```

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?

```
WITH t0 as (
        SELECT round(forest_percentage_region::numeric,2) as world_forest_percentage_1990
        FROM regions2
        WHERE region = 'World' and year = 1990
),

t1 as (
        SELECT year, region, MAX(forest_percentage_region) as max_1990
   FROM regions2
   WHERE year = 1990
   GROUP BY 1, 2
        ORDER BY 3 DESC
        LIMIT 1
```

```
),

t2 as (
        SELECT year, region, MIN(forest_percentage_region) as min_1990
   FROM regions2
   WHERE year = 1990
   GROUP BY 1, 2
        ORDER BY 3
        LIMIT 1
)

SELECT t0.world_forest_percentage_1990, t1.region as max_region, round(t1.max_1990::numeric,2) as
forest_percentage_max, t2.region as min_region, round(t2.min_1990::numeric,2) as forest_percentage_min
FROM t0, t1, t2
```

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

```
WITH regions90 as (
        SELECT *
        FROM regions2
        WHERE year = 1990
),

regions16 as (
        SELECT *
        FROM regions2
        WHERE year = 2016
)

SELECT regions90.region,                    round(regions90.forest_percentage_region::numeric,2) as
        forest_percentage_1990,          round(regions16.forest_percentage_region::numeric,2) as
        forest_percentage_2016
FROM regions90
JOIN regions16
        ON regions90.region = regions16.region
WHERE regions90.forest_percentage_region > regions16.forest_percentage_region
```

C. PART 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 countries16 as (
        SELECT *
```

```
        FROM forestation f
        WHERE year = 2016 AND f.forest_area_sq_km IS NOT NULL and f.country_name != 'World'
),

countries90 as (
        SELECT *
        FROM forestation f
        WHERE year = 1990 AND f.forest_area_sq_km IS NOT NULL AND f.country_name != 'World'
)

SELECT
        countries16.country_name, countries90.country_name, r.region,
   countries90.forest_area_sq_km as forest_area_1990,
   countries16.forest_area_sq_km as forest_area_2016,
   (countries90.forest_area_sq_km - countries16.forest_area_sq_km) AS total_difference
FROM countries16
JOIN countries90
        ON countries16.country_name = countries90.country_name
   AND (countries16.forest_area_sq_km IS NOT NULL
   AND countries90.forest_area_sq_km IS NOT NULL)
JOIN regions r
        ON r.country_code = countries16.country_code
ORDER BY 6 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 countries16 as (
        SELECT *
        FROM forestation f
        WHERE year = 2016 AND f.forest_area_sq_km IS NOT NULL and f.country_name != 'World'
),

countries90 as (
        SELECT *
        FROM forestation f
        WHERE year = 1990 AND f.forest_area_sq_km IS NOT NULL AND f.country_name != 'World'
)

SELECT
        countries16.country_name, countries90.country_name, r.region,
   countries90.forest_area_sq_km as forest_area_1990,
   countries16.forest_area_sq_km as forest_area_2016,
   (countries90.forest_area_sq_km - countries16.forest_area_sq_km) AS total_difference,
   ABS(ROUND(CAST(((countries16.forest_area_sq_km-
countries90.forest_area_sq_km)/countries90.forest_area_sq_km*100) AS NUMERIC),2)) AS percentage_difference
FROM countries16
JOIN countries90
        ON countries16.country_name = countries90.country_name
   AND (countries16.forest_area_sq_km IS NOT NULL
   AND countries90.forest_area_sq_km IS NOT NULL)
JOIN regions r
        ON r.country_code = countries16.country_code
```

```
ORDER BY
	ROUND(CAST(((countries16.forest_area_sq_km-
countries90.forest_area_sq_km)/countries90.forest_area_sq_km*100) AS NUMERIC),2)
LIMIT 5;
```

SUCESS STORIES PARAGRAPH 1 - Which 5 countries saw the largest amount INCREASED in forest area from 1990 to 2016? What was the difference in forest area for each?

```
WITH countries16 as (
	SELECT *
	FROM forestation f
	WHERE year = 2016 AND f.forest_area_sq_km IS NOT NULL and f.country_name != 'World'
),

countries90 as (
	SELECT *
	FROM forestation f
	WHERE year = 1990 AND f.forest_area_sq_km IS NOT NULL AND f.country_name != 'World'
)

SELECT
	countries16.country_name, countries90.country_name, r.region,
   countries90.forest_area_sq_km as forest_area_1990,
   countries16.forest_area_sq_km as forest_area_2016,
   (countries90.forest_area_sq_km - countries16.forest_area_sq_km) AS total_difference
FROM countries16
JOIN countries90
	ON countries16.country_name = countries90.country_name
   AND (countries16.forest_area_sq_km IS NOT NULL
   AND countries90.forest_area_sq_km IS NOT NULL)
JOIN regions r
	ON r.country_code = countries16.country_code
ORDER BY 6
LIMIT 5
```

SUCESS STORIES PARAGRAPH 2 - Which 5 countries saw the largest percent INCREASE in forest area from 1990 to 2016? What was the percent change to 2 decimal places for each?

```
WITH countries16 as (
	SELECT *
	FROM forestation f
	WHERE year = 2016 AND f.forest_area_sq_km IS NOT NULL and f.country_name != 'World'
),

countries90 as (
	SELECT *
	FROM forestation f
	WHERE year = 1990 AND f.forest_area_sq_km IS NOT NULL AND f.country_name !=
'World'
)

SELECT
```

```
        countries16.country_name, countries90.country_name, r.region,
   countries90.forest_area_sq_km as forest_area_1990,
   countries16.forest_area_sq_km as forest_area_2016,
   (countries90.forest_area_sq_km - countries16.forest_area_sq_km) AS total_difference,
   ABS(ROUND(CAST(((countries16.forest_area_sq_km-
countries90.forest_area_sq_km)/countries90.forest_area_sq_km*100) AS NUMERIC),2)) AS
percentage_difference
FROM countries16
JOIN countries90
        ON countries16.country_name = countries90.country_name
   AND (countries16.forest_area_sq_km IS NOT NULL
   AND countries90.forest_area_sq_km IS NOT NULL)
JOIN regions r
        ON r.country_code = countries16.country_code
ORDER BY
        ROUND(CAST(((countries16.forest_area_sq_km-
countries90.forest_area_sq_km)/countries90.forest_area_sq_km*100) AS NUMERIC),2) DESC
LIMIT 5;
```

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

```
WITH countries AS (
        SELECT
        f.country_code, f.country_name, f.year, f.forest_area_sqkm,
        l.total_area_sq_mi * 2.59 as total_area_sq_km,
        (f.forest_area_sqkm/(l.total_area_sq_mi*2.59))*100 AS fa_percentage
        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
),

quartiles as (
        SELECT c.country_code, c.country_name, c.year, c.fa_percentage,
                CASE WHEN c.fa_percentage >= 75 THEN 4
                WHEN c.fa_percentage < 75 AND c.fa_percentage >= 50 THEN 3
                WHEN c.fa_percentage < 50 AND c.fa_percentage >= 25 THEN 2
            ELSE 1 END AS quartile_group
        FROM countries c
        ORDER BY 5 DESC
)
```

```
SELECT quartiles.quartile_group, COUNT(quartiles.quartile_group)
FROM quartiles
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 countries AS (
            SELECT
            f.country_code, f.country_name, f.year, f.forest_area_sqkm,
            l.total_area_sq_mi * 2.59 as total_area_sq_km,
            (f.forest_area_sqkm/(l.total_area_sq_mi*2.59))*100 AS fa_percentage
            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
),

quartiles as (
            SELECT c.country_code, c.country_name, c.year, c.fa_percentage,
                        CASE WHEN c.fa_percentage >= 75 THEN 4
                        WHEN c.fa_percentage < 75 AND c.fa_percentage >= 50 THEN 3
                        WHEN c.fa_percentage < 50 AND c.fa_percentage >= 25 THEN 2
                    ELSE 1 END AS quartile_group
            FROM countries c
            ORDER BY 5 DESC
)

SELECT c.country_name, r.region, round(c.fa_percentage::numeric,2) as fa_percentage
FROM countries c
JOIN regions r
    ON r.country_code = c.country_code
WHERE year = 2016 AND fa_percentage >= 75
ORDER BY 3 DESC
```

E - How many countries had a percent forestation higher than the United States in 2016?

```
WITH countries as (
SELECT country_name, forest_percentage, year
FROM forestation
WHERE country_name != 'United States' AND country_name != 'World'
            AND year = 2016
ORDER BY 2 DESC
),

usa AS (
            SELECT country_name, forest_percentage, year
            FROM forestation
```

```
        WHERE country_name = 'United States'
                AND year = 2016
        ORDER BY 2 DESC
)

SELECT COUNT(c.country_name) as total_countries
FROM countries c
JOIN usa
        ON c.year = usa.year
WHERE c.forest_percentage > usa.forest_percentage
```