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 3.20824258980244%.

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).

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%
Middle East & North Africa	1.78%	2.07%
North America	35.65%	36.04%
East Asia & Pacific	25.78%	26.36%
Europe & Central Asia	37.28%	38.04%
South Asia	16.51%	17.51%

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 527229.062. 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 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	282193.9844
Myanmar	East Asia & Pacific	107234.0039
Nigeria	Sub-Saharan Africa	106506.00098
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%	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	98.2576939676578
Micronesia, Fed. Sts.	East Asia & Pacific	91.8572390715248
Gabon	Sub-Saharan Africa	90.0376418700565
Seychelles	Sub-Saharan Africa	88.4111367385789
Palau	East Asia & Pacific	87.6068085491204
American Samoa	East Asia & Pacific	87.5000875000875
Guyana	Latin America & Caribbean	83.9014489110682
Lao PDR	East Asia & Pacific	82.1082317640861
Solomon Islands	East Asia & Pacific	77.8635177945066

4. RECOMMENDATIONS

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

- *What have you learned from the World Bank data?*
    - *The decline in global forest area from 1990 to 2016 underlines the substantial loss of forests in Latin America and the Caribbean, as well as Sub-Saharan Africa. It's worth noting that despite this concerning trend, many regions around the world have shown positive developments in their forested areas. Notably, Europe and Central Asia, North America, East Asia and the Pacific, South Asia, and the Middle East and North Africa have all demonstrated an increase in both forested area and forest area percentage. This encouraging trend reflects their commitment to forest conservation and sustainable growth.*

- *Which countries should we focus on over others?*
    - *Evidently, a strategic focus on countries that have witnessed the most substantial reductions in forested areas, both in terms of percentage and total amount, is the most prudent approach. According to the findings in Table 3.2, Togo, Nigeria, Uganda, Mauritania, and Honduras are notable for the highest percentage decreases. Meanwhile, Table 3.1 identifies Brazil, Indonesia, Myanmar, Nigeria, and Tanzania as the countries with the most significant absolute decreases in forest area.*

5. APPENDIX: SQL Queries Used

Create a view:

```
CREATE VIEW forestation
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
        ,r.region
        ,r.income_group
        ,(f.forest_area_sqkm / (l.total_area_sq_mi * 2.59)) * 100 AS
perc_forest_area
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 r.country_code = f.country_code
```

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.

```sql
SELECT SUM(forest_area_sqkm)
FROM forestation
WHERE country_name = 'World'
        AND 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.”

```sql
SELECT SUM(forest_area_sqkm)
FROM forestation
WHERE country_name = 'World'
        AND year = 2016
```

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

```sql
SELECT A.forest_area_sqkm - B.forest_area_sqkm AS diff
FROM forestation A
INNER JOIN forestation B ON A.country_code = B.country_code
WHERE A.country_name = 'World'
        AND A.year = 1990
        AND B.year = 2016
```

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

```sql
SELECT (A.forest_area_sqkm - B.forest_area_sqkm) / A.forest_area_sqkm * 100
AS perc_diff
FROM forestation A
INNER JOIN forestation B ON A.country_code = B.country_code
WHERE A.country_name = 'World'
        AND A.year = 1990
        AND B.year = 2016
```

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
WITH V1
AS (
        SELECT SUM(forest_area_sqkm) AS forest_area_sqkm_1990
        FROM forestation
        WHERE country_name = 'World'
                AND year = 1990
        )
        ,V2
AS (
        SELECT SUM(forest_area_sqkm) AS forest_area_sqkm_2016
        FROM forestation
        WHERE country_name = 'World'
                AND year = 2016
        )
        ,V3
AS (
```

```
        SELECT (forest_area_sqkm_1990 - forest_area_sqkm_2016) AS diff
        FROM V1,V2
        )
SELECT country_name
        ,forest_area_sqkm
        ,ABS(forest_area_sqkm - (SELECT * FROM V3)) AS diff
FROM forestation
ORDER BY 3 ASC LIMIT 1
```

2. REGIONAL OUTLOOK

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 V1
AS (
        SELECT region
                ,year
                ,ROUND(CAST(SUM(forest_area_sqkm) / SUM(total_area_sqkm) * 100
AS NUMERIC), 2) AS perc
        FROM forestation
        GROUP BY 1,2
        HAVING year >= 1990
                AND year <= 2016
        )
SELECT *
FROM V1
WHERE region = 'World'
        AND year = 2016
WITH V1 AS (
                SELECT region
                        ,year
                        ,ROUND(CAST(SUM(forest_area_sqkm) /
SUM(total_area_sqkm) * 100 AS NUMERIC), 2) AS perc
                FROM forestation
                GROUP BY 1,2
                HAVING year >= 1990
                        AND year <= 2016
                )

SELECT *
FROM V1
WHERE year = 2016
ORDER BY perc DESC LIMIT 1
WITH V1 AS (
                SELECT region
                        ,year
                        ,ROUND(CAST(SUM(forest_area_sqkm) /
SUM(total_area_sqkm) * 100 AS NUMERIC), 2) AS perc
                FROM forestation
                GROUP BY 1,2
                HAVING year >= 1990
                        AND year <= 2016
                )

SELECT *
```

```
FROM V1
WHERE year = 2016
ORDER BY perc ASC 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?

```
WITH V1
AS (
        SELECT region
                ,year
                ,ROUND(CAST(SUM(forest_area_sqkm) / SUM(total_area_sqkm) * 100
AS NUMERIC), 2) AS perc
        FROM forestation
        GROUP BY 1,2
        HAVING year >= 1990
                AND year <= 2016
        )
SELECT *
FROM V1
WHERE region = 'World'
        AND year = 1990
WITH V1 AS (
                SELECT region
                        ,year
                        ,ROUND(CAST(SUM(forest_area_sqkm) /
SUM(total_area_sqkm) * 100 AS NUMERIC), 2) AS perc
                FROM forestation
                GROUP BY 1,2
                HAVING year >= 1990
                        AND year <= 2016
                )

SELECT *
FROM V1
WHERE year = 1990
ORDER BY perc DESC LIMIT 1
WITH V1 AS (
                SELECT region
                        ,year
                        ,ROUND(CAST(SUM(forest_area_sqkm) /
SUM(total_area_sqkm) * 100 AS NUMERIC), 2) AS perc
                FROM forestation
                GROUP BY 1,2
                HAVING year >= 1990
                        AND year <= 2016
                )

SELECT *
FROM V1
WHERE year = 1990
ORDER BY perc ASC
LIMIT 1
```

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

```
WITH V1
AS (
        SELECT region
                ,year
                ,ROUND(CAST(SUM(forest_area_sqkm) / SUM(total_area_sqkm) * 100
AS NUMERIC), 2) AS perc
        FROM forestation
        GROUP BY 1,2
        HAVING year >= 1990
                AND year <= 2016
        )
        ,V2
AS (
        SELECT *
        FROM V1
        WHERE year = 1990
        )
        ,V3
AS (
        SELECT *
        FROM V1
        WHERE year = 2016
        )
SELECT V2.region
        ,V2.perc AS perc_1990
        ,V3.perc AS perc_2016
        ,(V3.perc - V2.perc) AS diff
FROM V2
INNER JOIN V3 ON V2.region = V3.region
ORDER BY 4 ASC
```

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 V1
AS (
        SELECT country_name
                ,year
                ,region
                ,SUM(forest_area_sqkm) AS forest_area_sqkm_1990
        FROM forestation
        GROUP BY 1,2,3
        HAVING year = 1990
        )
        ,V2
AS (
        SELECT country_name
                ,year
                ,region
```

```
                ,SUM(forest_area_sqkm) AS forest_area_sqkm_2016
        FROM forestation
        GROUP BY 1,2,3
        HAVING year = 2016
        )
SELECT V1.country_name
        ,V1.region
        ,V1.forest_area_sqkm_1990
        ,V2.forest_area_sqkm_2016
        ,(V1.forest_area_sqkm_1990 - V2.forest_area_sqkm_2016) AS diff
        ,(V1.forest_area_sqkm_1990 - V2.forest_area_sqkm_2016) /
V1.forest_area_sqkm_1990 * 100 AS perc
FROM V1
INNER JOIN V2 ON V1.country_name = V2.country_name
WHERE V1.forest_area_sqkm_1990 IS NOT NULL
        AND V2.forest_area_sqkm_2016 IS NOT NULL
ORDER BY 5 DESC
```

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 V1
AS (
        SELECT country_name
                ,year
                ,region
                ,SUM(forest_area_sqkm) AS forest_area_sqkm_1990
        FROM forestation
        GROUP BY 1,2,3
        HAVING year = 1990
        )
        ,V2
AS (
        SELECT country_name
                ,year
                ,region
                ,SUM(forest_area_sqkm) AS forest_area_sqkm_2016
        FROM forestation
        GROUP BY 1,2,3
        HAVING year = 2016
        )
SELECT V1.country_name
        ,V1.region
        ,V1.forest_area_sqkm_1990
        ,V2.forest_area_sqkm_2016
        ,(V1.forest_area_sqkm_1990 - V2.forest_area_sqkm_2016) AS diff
        ,ROUND(CAST((V1.forest_area_sqkm_1990 - V2.forest_area_sqkm_2016) /
V1.forest_area_sqkm_1990 * 100 AS NUMERIC), 2) AS perc
FROM V1
INNER JOIN V2 ON V1.country_name = V2.country_name
WHERE V1.forest_area_sqkm_1990 IS NOT NULL
        AND V2.forest_area_sqkm_2016 IS NOT NULL
ORDER BY 6 DESC
```

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

```
WITH V1
AS (
        SELECT country_name
                ,CASE
                        WHEN perc_forest_area < 25
                                THEN '0-25%'
                        WHEN perc_forest_area >= 25
                                AND perc_forest_area < 50
                                THEN '25-50%'
                        WHEN perc_forest_area >= 50
                                AND perc_forest_area < 75
                                THEN '50-75%'
                        ELSE '75-100%'
                        END AS percentiles
        FROM forestation
        WHERE year = 2016
        )
SELECT percentiles
        ,COUNT(*)
FROM V1
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 V1
AS (
        SELECT country_name
                ,region
                ,perc_forest_area
                ,CASE
                        WHEN perc_forest_area < 25
                                THEN '0-25%'
                        WHEN perc_forest_area >= 25
                                AND perc_forest_area < 50
                                THEN '25-50%'
                        WHEN perc_forest_area >= 50
                                AND perc_forest_area < 75
                                THEN '50-75%'
                        ELSE '75-100%'
                        END AS percentiles
        FROM forestation
        WHERE year = 2016
                AND perc_forest_area IS NOT NULL
        )
SELECT *
FROM V1
WHERE percentiles = '75-100%'
ORDER BY 3 DESC
```

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

```
WITH V1
AS (
        SELECT perc_forest_area
        FROM forestation
        WHERE year = 2016
                AND country_name = 'United States'
        )
SELECT COUNT(*)
FROM forestation
WHERE perc_forest_area > (SELECT * FROM V1)
```