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 km²___ in 1990. As of 2016, the most recent year for which data was available, that number had fallen to _____39958245.9 km²____, a loss of _____1324449 km²__, or _____3.2082_____%.

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 km²_____).

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 _____41.16_%, 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
Sub-Saharan Africa	30.67	28.79
World	32.42	31.38
Latin America & Caribbean	51.03	46.16

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 527229.062 km^2. 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 km^2, 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 (should it be 5?) 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 (km^2)
Brazil	Latin America & Caribbean	541510
China	East Asia & Pacific	527229.06
Indonesia	East Asia & Pacific	282193.98
Myanmar	East Asia & Pacific	107234
Nigeria	Sub-Saharan Africa	106506

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 (% decrease)
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 (0 - 25%)	85
2 (25 - 50%)	73
3 (50 - 75%)	38
4 (75 - 100%)	9

The largest number of countries in 2016 were found in the __1st__ 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

Approximately 2,516 countries had a higher percentage of forest than the United States in 2016.

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 data shows the pressing dangers of deforestation and the alarming rate at which the world's forests are being depleted. Looking at the data presents a dilemma on where to put our efforts; do we tackle the issue of where the amount of deforestation is largest in means of square kilometers, like in the case of China, or do we tackle the issue where deforestation is largest in means of percentage of forest lost? The data suggests that the latter is the better option for remedying the issue.

- *Which countries should we focus on over others?*
    - The countries that we should focus our efforts on most are those listed in Table 3.2, especially those in the Sub-Saharan Africa regions. The most concerning of the countries is Togo, which has lost over ¾ of its forests in the span from 1990 to 2016. Our efforts should be aimed mainly at these smaller countries that have been more heavily affected, but we should also keep an eye on the sources of the largest deforestation, namely Brazil and China. As seen in the case of Iceland, deforestation doesn't have to be an inevitable part of the passage of time.

5. APPENDIX

```
/* 1. Create View called "forestation" that joins all three tables

   2: The forest_area and land_area tables join on both country_code
      AND year.

   3: The regions table joins these based on only country_code.

   4: In the 'forestation' View, include the following: All of the
      columns of the origin tables A new column that provides the percent
      of the land area that is designated as forest

   5: Keep in mind that the column forest_area_sqkm in the forest_area
      table and the land_area_sqmi in the land_area table are in different
      units (square kilometers and square miles, respectively), so an
      adjustment will need to be made in the calculation you write
      (1 sq mi = 2.59 sq km). */


CREATE VIEW forestation AS
SELECT  f.country_code AS country_code,
        f.country_name AS country_name,
        f.year AS year,
        r.region AS region,
        r.income_group AS income_group,
        f.forest_area_sqkm AS forest_area_sqkm,
        l.total_area_sq_mi AS total_area_sq_mi,
        ROUND((SUM(forest_area_sqkm) / (SUM(total_area_sq_mi)*2.59) *
100):: numeric, 2) AS percent_forest
FROM forest_area f
FULL OUTER JOIN land_area l
ON f.country_code = l.country_code AND f.year = l.year
FULL OUTER JOIN regions r
ON f.country_code = r.country_code
GROUP BY 1,2,3,4,5,6,7
ORDER BY 3,4,5,6,7,8
```

```sql
                    --Part 1: Global Situation Queries --
/* a: What was the total forest area (in sq km) of the world in 1990?
      Keep in mind that you can use the country record denoted as "World"
      in the regions table */

SELECT  year,
        SUM(forest_area_sqkm)
FROM forestation
WHERE year = '1990' and country_name = 'World'
GROUP BY 1
ORDER BY 2


/* 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,
        SUM(forest_area_sqkm)
FROM forestation
WHERE year = '2016' and country_name = 'World'
GROUP BY 1
ORDER BY 2



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

WITH total_forest_1990 AS (
    SELECT  year,
            country_name,
            SUM(forest_area_sqkm) AS total_forest_area
    FROM forestation
    WHERE year = '1990' and country_name = 'World'
```

```
    GROUP BY 1,2
    ORDER BY 3
),
    total_forest_2016 AS (
    SELECT  year,
            country_name,
            SUM(forest_area_sqkm) AS total_forest_area
    FROM forestation
    WHERE year = '2016' and country_name = 'World'
    GROUP BY 1,2
    ORDER BY 3
    )
SELECT (x.total_forest_area - y.total_forest_area) AS global_change
FROM total_forest_1990 AS x
JOIN total_forest_2016 AS y
ON x.country_name = y.country_name



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

WITH total_forest_1990 AS (
    SELECT  year,
            country_name,
            SUM(forest_area_sqkm) AS total_forest_area
    FROM forestation
    WHERE year = '1990' and country_name = 'World'
    GROUP BY 1,2
    ORDER BY 3
),
    total_forest_2016 AS (
    SELECT  year,
            country_name,
            SUM(forest_area_sqkm) AS total_forest_area
    FROM forestation
    WHERE year = '2016' and country_name = 'World'
```

```
    GROUP BY 1,2
    ORDER BY 3
    )
SELECT ((x.total_forest_area - y.total_forest_area)/x.total_forest_area *
100) AS percent_global_change
FROM total_forest_1990 AS x
JOIN total_forest_2016 AS y
ON x.country_name = y.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,
       (total_area_sq_mi * 2.59) AS total_area_sqkm
FROM forestation
WHERE year = 2016 AND (total_area_sq_mi * 2.59) < '1324449'
ORDER BY total_area_sqkm DESC
LIMIT 1;




           -- 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). */

CREATE TABLE regional_outlook AS
           SELECT region,
                  year,
                  ((SUM(forest_area_sqkm) / (SUM(total_area_sq_mi) *
2.59)) * 100) AS percent_forest
FROM forestation
WHERE year = '1990' OR year = '2016'
```

```sql
GROUP BY 1, 2
ORDER BY 1, 2, 3



/* 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? */

-- World--
SELECT region,
       year,
       CAST((percent_forest) AS DECIMAL (5,2))
FROM forestation
WHERE year = '2016' AND region = 'World'
GROUP BY 1, 2, 3
ORDER BY 3

-- HIGHEST --
SELECT region,
       year,
       CAST((percent_forest) AS DECIMAL (5,2))
FROM regional_outlook
WHERE year = '2016' AND percent_forest IS NOT NULL
GROUP BY 1, 2, 3
ORDER BY 3 DESC
LIMIT 1

--LOWEST
SELECT region,
       year,
       CAST((percent_forest) AS DECIMAL (5,2))
FROM regional_outlook
WHERE year = '2016' AND percent_forest IS NOT NULL
GROUP BY 1, 2, 3
ORDER BY 3
LIMIT 1
```

```sql
/* 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? */

-- World--
SELECT region,
       year,
       CAST((percent_forest) AS DECIMAL (5,2))
FROM regional_outlook
WHERE year = '1990' AND region = 'World'
GROUP BY 1, 2, 3
ORDER BY 3

-- HIGHEST --
SELECT region,
       year,
       CAST((percent_forest) AS DECIMAL (5,2))
FROM regional_outlook
WHERE year = '1990' AND percent_forest IS NOT NULL
GROUP BY 1, 2, 3
ORDER BY 3 DESC
LIMIT 1

--LOWEST
SELECT region,
       year,
       CAST((percent_forest) AS DECIMAL (5,2))
FROM regional_outlook
WHERE year = '1990' AND percent_forest IS NOT NULL
GROUP BY 1, 2, 3
ORDER BY 3
LIMIT 1



/* c:  Based on the table you created, which regions of the world
```

```
       DECREASED in forest area from 1990 to 2016? */

WITH total_forest_1990 AS (
    SELECT  region,
            CAST((percent_forest) AS DECIMAL (5,2))
    FROM regional_outlook
    WHERE year = '1990'
),

     total_forest_2016 AS (
        SELECT  region,
                CAST((percent_forest) AS DECIMAL (5,2))
        FROM regional_outlook
        WHERE year = '2016'
     )
SELECT  x.region,
        x.percent_forest AS percent_1990,
        y.percent_forest AS percent_2016,
        CASE WHEN (x.percent_forest - y.percent_forest) > 0 THEN
'DECREASE'
        ELSE 'INCREASE' END
FROM total_forest_1990 AS x
JOIN total_forest_2016 AS y
ON x.region = y.region
WHERE (x.percent_forest - y.percent_forest) > 0
GROUP BY 1, 2, 3, 4
ORDER BY 2, 3




            -- Part 3: Country Level Detail --
/*  Unmarked Question:
    1: Find the two countries that had the largest increase in forest
       between 1990 and 2016*/
WITH  countries_1990 AS (
        SELECT  country_name,
```

```
                region,
                year,
                forest_area_sqkm
        FROM    forestation
        WHERE   year = '1990'
        GROUP BY 1, 2, 3, 4
        ORDER BY 4
),

      countries_2016 AS (
        SELECT  country_name,
                region,
                year,
                forest_area_sqkm
        FROM    forestation
        WHERE   year = '2016'
        GROUP BY 1, 2, 3, 4
        ORDER BY 4
    )
SELECT  x.country_name AS country_1990,
        x.region,
        (x.forest_area_sqkm - y.forest_area_sqkm) AS forest_difference

FROM countries_1990 AS x
JOIN countries_2016 AS y
ON x.country_name = y.country_name
WHERE (x.forest_area_sqkm - y.forest_area_sqkm) IS NOT NULL
  AND x.country_name != 'World' AND y.country_name != 'World'
GROUP BY 1, 2, 3
ORDER BY 3
LIMIT 2;


/*  Unmarked Question:
    2:  Find the percentage of change for each country, then find the
        country with the largest percentage: */
```

```sql
WITH  countries_1990 AS (
       SELECT  country_name,
               region,
               year,
               forest_area_sqkm
       FROM    forestation
       WHERE   year = '1990'
       GROUP BY 1, 2, 3, 4
       ORDER BY 4
),

     countries_2016 AS (
       SELECT  country_name,
               region,
               year,
               forest_area_sqkm
       FROM    forestation
       WHERE   year = '2016'
       GROUP BY 1, 2, 3, 4
       ORDER BY 4
    )
SELECT  x.country_name AS country_1990,
       x.region,
       ABS((x.forest_area_sqkm - y.forest_area_sqkm)) AS
forest_difference,
       ((x.forest_area_sqkm - y.forest_area_sqkm)/ x.forest_area_sqkm *
100) AS percentage_change
FROM countries_1990 AS x
JOIN countries_2016 AS y
ON x.country_name = y.country_name
WHERE (x.forest_area_sqkm - y.forest_area_sqkm) IS NOT NULL
  AND x.country_name != 'World' AND y.country_name != 'World'
GROUP BY 1, 2, 3, 4
ORDER BY 4
LIMIT 1;
```

```sql
/* 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  countries_1990 AS (
        SELECT  country_name,
                region,
                year,
                forest_area_sqkm
        FROM    forestation
        WHERE   year = '1990'
        GROUP BY 1, 2, 3, 4
        ORDER BY 4
),

      countries_2016 AS (
        SELECT  country_name,
                region,
                year,
                forest_area_sqkm
        FROM    forestation
        WHERE   year = '2016'
        GROUP BY 1, 2, 3, 4
        ORDER BY 4
    )
SELECT  x.country_name AS country_1990,
        x.region,
        ABS((x.forest_area_sqkm - y.forest_area_sqkm)) AS
forest_difference,
        ((x.forest_area_sqkm - y.forest_area_sqkm)/ x.forest_area_sqkm *
100) AS percentage_change
FROM countries_1990 AS x
JOIN countries_2016 AS y
ON x.country_name = y.country_name
WHERE (x.forest_area_sqkm - y.forest_area_sqkm) IS NOT NULL
```

```
  AND x.country_name != 'World' AND y.country_name != 'World'
GROUP BY 1, 2, 3, 4
ORDER BY 3, 4 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  countries_1990 AS (
        SELECT  country_name,
                region,
                year,
                forest_area_sqkm
        FROM    forestation
        WHERE   year = '1990'
        GROUP BY 1, 2, 3, 4
        ORDER BY 4
),

      countries_2016 AS (
        SELECT  country_name,
                region,
                year,
                forest_area_sqkm
        FROM    forestation
        WHERE   year = '2016'
        GROUP BY 1, 2, 3, 4
        ORDER BY 4
    )
SELECT  x.country_name AS country_1990,
        x.region,
```

```
       ABS((x.forest_area_sqkm - y.forest_area_sqkm)) AS
forest_difference,
       ((x.forest_area_sqkm - y.forest_area_sqkm)/ x.forest_area_sqkm *
100) AS percentage_change
FROM countries_1990 AS x
JOIN countries_2016 AS y
ON x.country_name = y.country_name
WHERE (x.forest_area_sqkm - y.forest_area_sqkm) IS NOT NULL
  AND x.country_name != 'World' AND y.country_name != 'World'
GROUP BY 1, 2, 3, 4
ORDER BY 4 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_2016 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 country_name != 'World' AND percent_forest IS
NOT NULL)
SELECT  quartile,
       COUNT(quartile) count
FROM countries_2016
GROUP BY 1
ORDER BY 2;
```

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

WITH countries_2016 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 country_name != 'World' AND percent_forest IS
NOT NULL)
SELECT  country_name,
        quartile,
        percent_forest
FROM countries_2016
WHERE quartile = 4
GROUP BY 1, 2, 3
ORDER BY 3 DESC;



/* 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')
```