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 km^2 in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 39958245.9 km^2, a loss of 1324449 km^2, or 3.21%.

The forest area lost over this time period is slightly more than the entire land area of Mongolia listed for the year 2016 (which is 1553560 km^2).

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
East Asia & Pacific	25.78	26.36
Europe & Central Asia	37.28	38.04
Latin America & Caribbean	51.03	46.16
Middle East & North Africa	1.78	2.07
North America	35.65	36.04
South Asia	16.51	17.51
Sub-Saharan Africa	30.67	28.79

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 527229.06 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 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 313.66% 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 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
Indonesia	East Asia & Pacific	-282194
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	-61.8
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
First	85
Second	73
Third	38
Fourth	9

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

4. RECOMMENDATIONS

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

- *What have you learned from the World Bank data?*
- *Which countries should we focus on over others?*

Based on analysis from the World Bank, there are signs of further shrinkage of global forest area. Most attention should be focused on countries where there has been the greatest percentage reduction in forest area. Those countries are Togo, Nigeria, Uganda, Mauritania and Honduras. Four out of five those countries are located in the Sub-Saharan Africa region and all have low to lower middle income. It is important to assess the reasons for the decline in forest cover in these regions. Whether this is due to sourcing of exotic wood, agriculture, fires or climate changes. A good example of a country that has increased its forest area is China. This case should be studied and recommendations made for low income countries.

5. APPENDIX: SQL queries used

```sql
-- Create a View by joining all three tables - forest_area, land_area and regions
CREATE VIEW forestation AS
SELECT fa.country_code,
  fa.country_name,
  r.region,
  r.income_group,
  fa.year,
  fa.forest_area_sqkm,
  -- convert land area to square kilimeters
  la.total_area_sq_mi * 2.59 AS total_area_sqkm,
  -- compute the percent of forest area
  (
    fa.forest_area_sqkm / (la.total_area_sq_mi * 2.59)
  ) * 100 AS forest_percent
FROM forest_area AS fa -- join land and forest tables
  JOIN land_area AS la ON fa.country_code = la.country_code
  AND fa.year = la.year -- join region to forest table
  JOIN regions AS r ON fa.country_code = r.country_code
ORDER BY country_code,
  year;
```

```sql
-- PART 1
-- GLOBAL SITUATION

-- Q1
/* 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 ROUND(forest_area_sqkm) AS forest_area_sqkm_1990
FROM forestation
WHERE year = 1990 AND country_name = 'World';

-- Q2
/* 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 ROUND(forest_area_sqkm) AS forest_area_sqkm_2016
FROM forestation
WHERE year = 2016 AND country_name = 'World';

-- Q3
/* What was the change (in sq km) in the forest area of the world from 1990 to
2016? */
WITH table1 AS(
  SELECT forest_area_sqkm AS forest_area_sqkm_1990
  FROM forestation
  WHERE year = 1990
    AND country_name = 'World'
),
table2 AS(
  SELECT forest_area_sqkm AS forest_area_sqkm_2016
  FROM forestation
  WHERE year = 2016
    AND country_name = 'World'
)
SELECT ROUND(table1.forest_area_sqkm_1990 - table2.forest_area_sqkm_2016) AS
forest_area_change
FROM table1,
  table2;
```

```sql
-- Q4
/* What was the percent change in forest area of the world between 1990 and 2016?
*/
WITH table1 AS(
  SELECT forest_area_sqkm AS forest_area_sqkm_1990
  FROM forestation
  WHERE year = 1990
    AND country_name = 'World'
),
table2 AS(
  SELECT forest_area_sqkm AS forest_area_sqkm_2016
  FROM forestation
  WHERE year = 2016
    AND country_name = 'World'
)
SELECT ROUND(
    (
      (
        -(
          table1.forest_area_sqkm_1990 - table2.forest_area_sqkm_2016
        ) / table1.forest_area_sqkm_1990
      ) * 100
    )::NUMERIC,
    2
  )::VARCHAR || '%' AS forest_area_change
FROM table1,
  table2;
```

```sql
-- Q5
/* 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 DISTINCT country_name,
  ROUND(total_area_sqkm::NUMERIC) AS total_area_sqkm
FROM forestation
WHERE total_area_sqkm >= (
    WITH table1 AS(
      SELECT forest_area_sqkm AS forest_area_sqkm_1990
      FROM forestation
      WHERE year = 1990
        AND country_name = 'World'
    ),
    table2 AS(
      SELECT forest_area_sqkm AS forest_area_sqkm_2016
      FROM forestation
      WHERE year = 2016
        AND country_name = 'World'
    )
    SELECT table1.forest_area_sqkm_1990 - table2.forest_area_sqkm_2016 AS
forest_area_change
    FROM table1,
      table2
  )
ORDER BY total_area_sqkm
LIMIT 1;

-- PART 2
-- REGIONAL OUTLOOK

-- Q1
/* 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 sub.*,
  ROUND(
    ((sub.forest_area / sub.land_area) * 100)::NUMERIC,
    2
  ) AS forest_percent
FROM(SELECT region,
      SUM(forest_area_sqkm) AS forest_area,
      SUM(total_area_sqkm) AS land_area
    FROM forestation
    GROUP BY region,
      year
```

```
    HAVING year = 2016) AS sub
ORDER BY region;
-- Q2
/* 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 sub.*,
  ROUND(
    ((sub.forest_area / sub.land_area) * 100)::NUMERIC,
    2
  ) AS forest_percent
FROM(
    SELECT region,
      SUM(forest_area_sqkm) AS forest_area,
      SUM(total_area_sqkm) AS land_area
    FROM forestation
    GROUP BY region,
      year
    HAVING year = 1990
  ) AS sub
ORDER BY region;
```

```sql
-- Q3
/* Based on the table you created,
 which regions of the world DECREASED in forest area from 1990 to 2016? */
WITH t1 AS (
  SELECT sub.*,
    ROUND(
      ((sub.forest_area / sub.land_area) * 100)::NUMERIC,
      2
    ) AS forest_percent
  FROM(
      SELECT region,
        SUM(forest_area_sqkm) AS forest_area,
        SUM(total_area_sqkm) AS land_area
      FROM forestation
      GROUP BY region,
        year
      HAVING year = 2016
    ) AS sub
  ORDER BY forest_percent
),
t2 AS (
  SELECT sub.*,
    ROUND(
      ((sub.forest_area / sub.land_area) * 100)::NUMERIC,
      2
    ) AS forest_percent
  FROM(
      SELECT region,
        SUM(forest_area_sqkm) AS forest_area,
        SUM(total_area_sqkm) AS land_area
      FROM forestation
      GROUP BY region,
        year
      HAVING year = 1990
    ) AS sub
  ORDER BY forest_percent
)
SELECT t1.region,
  t1.forest_percent - t2.forest_percent AS change_prc
FROM t1
  JOIN t2 ON t1.region = t2.region
  AND t1.forest_percent < t2.forest_percent
ORDER BY change_prc;
```

```sql
-- PART 3
-- COUNTRY-LEVEL DETAIL

-- Q1
/* 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 t1 AS (
  SELECT country_code,
    country_name,
    region,
    forest_area_sqkm
  FROM forestation
  WHERE year = 1990
),
t2 AS (
  SELECT country_code,
    country_name,
    forest_area_sqkm
  FROM forestation
  WHERE year = 2016
)
SELECT t1.country_name,
  t1.region,
  t1.forest_area_sqkm AS forest_area_1990,
  t2.forest_area_sqkm AS forest_area_2016,
  ROUND(
    (t2.forest_area_sqkm - t1.forest_area_sqkm)::NUMERIC,
    2
  ) AS change
FROM t1
  JOIN t2 ON t1.country_code = t2.country_code
WHERE t1.country_name NOT LIKE 'World'
ORDER BY change
LIMIT 5;
```

```sql
-- Q2
/* 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_code,
    country_name,
    region,
    forest_area_sqkm
  FROM forestation
  WHERE year = 1990
),
t2 AS (
  SELECT country_code,
    country_name,
    forest_area_sqkm
  FROM forestation
  WHERE year = 2016
)
SELECT t1.country_name,
  t1.region,
  t1.forest_area_sqkm AS forest_area_1990,
  t2.forest_area_sqkm AS forest_area_2016,
  ROUND(
    -((1 -(t2.forest_area_sqkm / t1.forest_area_sqkm)) * 100)::NUMERIC, 2) AS
change_prc
FROM t1
  JOIN t2 ON t1.country_code = t2.country_code
  AND t2.forest_area_sqkm < t1.forest_area_sqkm
WHERE t1.country_name NOT LIKE 'World'
ORDER BY change_prc
LIMIT 5;
```

```sql
-- Q3
/* If countries were grouped by percent forestation in quartiles,
 which group had the most countries in it in 2016? */
WITH sub AS (
  SELECT country_name,
    CASE
      WHEN forest_percent < 25 THEN '0-25%'
      WHEN forest_percent >= 25
      AND forest_percent < 50 THEN '25-50%'
      WHEN forest_percent >= 50
      AND forest_percent < 75 THEN '50-75%'
      ELSE '75-100%'
    END AS quartile
  FROM forestation
  WHERE year = 2016
    AND forest_percent IS NOT NULL
)
SELECT DISTINCT quartile,
  (COUNT(country_name) OVER (PARTITION BY quartile)) AS count
FROM sub
ORDER BY quartile;

-- Q4
/* List all of the countries that were in the 4th quartile (percent forest > 75%)
in 2016. */
WITH sub AS (
  SELECT country_name,
    CASE
      WHEN forest_percent < 25 THEN '0-25%'
      WHEN forest_percent >= 25
      AND forest_percent < 50 THEN '25-50%'
      WHEN forest_percent >= 50
      AND forest_percent < 75 THEN '50-75%'
      ELSE '75-100%'
    END AS quartile
  FROM forestation
  WHERE year = 2016
    AND forest_percent IS NOT NULL
)
SELECT country_name,
  quartile
FROM sub
WHERE quartile = '75-100%';
```

```sql
-- Q5
/* How many countries had a percent forestation higher than the United States in
2016? */
SELECT COUNT(*) AS count
FROM(
    SELECT DISTINCT country_name
    FROM forestation
    WHERE forest_percent > (
        SELECT forest_percent
        FROM forestation
        WHERE (country_name = 'United States')
          AND year = 2016
      )
    ORDER BY country_name
  ) AS sub;
```