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 3.21%.

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 1279999.98 **km²**).

2. **REGIONAL OUTLOOK**

In 2016, the percentage 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 the Caribbean, with 46.16%, and the region with the lowest relative forestation was the Middle East and North Africa, with 2.07% forestation.

In 1990, the percentage 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 the Caribbean, with 51.03%, and the region with the lowest relative forestation was the 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
Latin America and 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 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 527,229.06 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 79,200 km^2, much lower than the figure for China.

China and the U.S 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 **km²**
Indonesia	East Asia & Pacific	282193.98 **km²**
Myanmar	East Asia & Pacific	107234 **km²**

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%

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 and 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	72

50 - 75	38
75 - 100	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

4. RECOMMENDATIONS

Write out a set of recommendations as an analyst on the ForestQuery team.
- *What have you learned from the World Bank data?*
- ➔ *At a global level, the total area of world forest has had a decrease of 3.21% between the years 1990 and 2016.*
- ➔ *At the regional level, Latin America and Caribbean dropped from 51.03% to 46.16% and Sub-Saharan Africa dropped from 30.67% to 28.79%, but the other regions have increased their forest area.*
- ➔ *China has increased in forest area by 527,229.06* km² *followed by the U.S by 79,200* km²*.*
- ➔ *Brazil, Indonesia and Myanmar are the top countries that have decreased their forest area in* km²*.*
- ➔ *Regarding the quartiles, there are 9 countries that have a high percentage of forest loss.*

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

- ➔ Forests help to minimize the effects produced by climate change, therefore I believe that a global approach is essential, to create awareness to avoid losses due to fires, to exploit forests in a more sustainable way, for example. Continue strengthening those regions and countries that have shown an increase in their forest area, see how they have done it and see if it can be implemented in other countries.

5. APPENDIX: SQL Queries Used

Project introduction (Create a View):

```sql
DROP VIEW IF EXISTS forestation;

CREATE VIEW forestation

AS

SELECT f.country_code,
       f.country_name,
       f.year,
       f.forest_area_sqkm,
       r.region,
       r.income_group,
       ( l.total_area_sq_mi * 2.59 ) AS total_area_sqkm,
       Round(( ( f.forest_area_sqkm / ( l.total_area_sq_mi * 2.59 ) ) * 100
)::numeric, 2) AS percent_forest
 FROM    forest_area f
       JOIN land_area l
         ON f.country_code = l.country_code
            AND f.year = l.year
       JOIN regions r
         ON r.country_code = f.country_code;
```

1. Global Situation

A.
```sql
SELECT Sum(forest_area_sqkm) AS total_forest_area

FROM forestation

WHERE year = 1990

    AND country_name = 'World';
```

B.
```sql
SELECT Sum(forest_area_sqkm) AS total_forest_area

FROM forestation

WHERE year = 2016

    AND country_name = 'World';
```

C.
```sql
SELECT (

        SELECT Sum(forest_area_sqkm) AS total_forest_area

        FROM forestation
```

```sql
        WHERE year = 1990
            AND country_name = 'World'
    ) - (
        SELECT Sum(forest_area_sqkm) AS total_forest_area
        FROM forestation
        WHERE year = 2016
            AND country_name = 'World'
    ) AS change_forest_area;
```

D.
```sql
WITH total_forest_area_1990 AS (

SELECT Sum(forest_area_sqkm) AS total_forest_area
        FROM    forestation
        WHERE   year = 1990
                AND country_name = 'World'
)

, total_forest_area_2016 AS (

SELECT Sum(forest_area_sqkm) AS total_forest_area
        FROM    forestation
        WHERE   year = 2016
                AND country_name = 'World')


SELECT ROUND(((((ELECT * FROM total_forest_area_1990) - (SELECT * FROM
total_forest_area_2016)) / (SELECT * FROM total_forest_area_1990)) * 100)::numeric, 2)
AS diff_percent;
```

E.
```sql
WITH total_forest_area_1990 AS (

SELECT Sum(forest_area_sqkm) AS total_forest_area
        FROM    forestation
        WHERE   year = 1990
                AND country_name = 'World'
)

, total_forest_area_2016 AS (

SELECT Sum(forest_area_sqkm) AS total_forest_area
        FROM    forestation
        WHERE   year = 2016
                AND country_name = 'World')


SELECT country_name,

total_area_sqkm,
```

```sql
ROUND(((abs(((SELECT * FROM total_forest_area_1990) - (SELECT * FROM
total_forest_area_2016)) - total_area_sqkm)/total_area_sqkm)*100)::numeric, 2) AS err
FROM forestation
WHERE year = '2016'
ORDER BY err
LIMIT 1;
```

2. Regional Outlook

```sql
--all question
WITH total_percent_forest AS (
SELECT region,
year,
ROUND(((sum(forest_area_sqkm) / sum(total_area_sqkm))*100)::numeric, 2) AS
percent_forest
FROM forestation
WHERE year IN (1990, 2016)
GROUP BY 1, 2)


SELECT region,
sum(case when year = 1990 then percent_forest else 0 end) as percent_forest_1990,
sum(case when year = 2016 then percent_forest else 0 end) as percent_forest_2016,
sum(case when year = 2016 then percent_forest else percent_forest*-1 end) as
decrease_forest
FROM total_percent_forest
GROUP BY 1
ORDER BY 4;
```

3. Country-level detail

Success Stories

```sql
SS1.
WITH sf_1990 AS
(SELECT country_name, region, forest_area_sqkm
FROM forestation
WHERE year = 1990
AND forest_area_sqkm IS NOT NULL
AND region != 'World'),
sf_2016 AS
(SELECT country_name, region, forest_area_sqkm
```

```sql
FROM forestation
WHERE year = 2016
AND forest_area_sqkm IS NOT NULL
AND region != 'World')
SELECT sf_1990.country_name, sf_1990.region, (sf_2016.forest_area_sqkm -
sf_1990.forest_area_sqkm) AS diff_forest_area
FROM sf_2016
JOIN sf_1990
USING (country_name)
WHERE sf_2016.forest_area_sqkm > sf_1990.forest_area_sqkm
ORDER BY 3 desc
LIMIT 5;
```

SS2.

```sql
SELECT sub1.country_name, sub2.region, ROUND(CAST( (((sub1.forest_area_sqkm -
sub2.forest_area_sqkm) * 100) / (sub2.forest_area_sqkm)) AS NUMERIC), 2) AS
diff_percent
FROM
  (SELECT country_name, region, forest_area_sqkm
   FROM forestation
   WHERE year = 2016) sub1
   JOIN
   (SELECT country_name, region, forest_area_sqkm
   FROM forestation
   WHERE year = 1990) sub2
ON sub1.country_name = sub2.country_name
WHERE (sub1.forest_area_sqkm, sub2.forest_area_sqkm) IS NOT NULL
ORDER BY 3 DESC
LIMIT 5;
```

Largest Concerns

A.

```sql
WITH sf_1990 AS
(SELECT country_name, region, forest_area_sqkm
FROM forestation
WHERE year = 1990
AND forest_area_sqkm IS NOT NULL
AND region != 'World'),
```

```sql
    sf_2016 AS
    (SELECT country_name, region, forest_area_sqkm
    FROM forestation
    WHERE year = 2016
    AND forest_area_sqkm IS NOT NULL
    AND region != 'World')

    SELECT sf_1990.country_name, sf_1990.region, (sf_1990.forest_area_sqkm -
    sf_2016.forest_area_sqkm) AS diff_forest_area
    FROM sf_1990
    JOIN sf_2016
    USING (country_name)
    WHERE sf_1990.forest_area_sqkm > sf_2016.forest_area_sqkm
    ORDER BY 3 DESC
    LIMIT 5;


B.
    WITH pf_1990 AS
    (SELECT country_name, region, forest_area_sqkm
    FROM forestation
    WHERE year = 1990
    AND forest_area_sqkm IS NOT NULL
    AND region != 'World'),

    pf_2016 AS
    (SELECT country_name, region, forest_area_sqkm
    FROM forestation
    WHERE year = 2016
    AND forest_area_sqkm IS NOT NULL
    AND region != 'World')

    SELECT pf_2016.country_name, pf_2016.region, ROUND(CAST( (((pf_1990.forest_area_sqkm -
    pf_2016.forest_area_sqkm) * 100) / (pf_1990.forest_area_sqkm)) AS NUMERIC), 2) AS
    percent_diff
    from pf_1990
    JOIN pf_2016
    USING (country_name)
    WHERE pf_1990.forest_area_sqkm > pf_2016.forest_area_sqkm
    ORDER BY 3 DESC
```

```sql
LIMIT 5;
```

Quartiles

C.

```sql
WITH range_quart AS
(SELECT country_name, region,
(CASE
WHEN percent_forest BETWEEN 75 AND 100 THEN 'Fourth'
WHEN percent_forest <= 75 AND percent_forest > 50 THEN 'Third'
WHEN percent_forest <= 50 AND percent_forest > 25 THEN 'Second'
ELSE 'First' END) AS range_quartiles
FROM forestation
WHERE year = 2016
AND region != 'World'
AND percent_forest IS NOT NULL)

SELECT COUNT(country_name) AS most_countries, range_quartiles
FROM range_quart
GROUP BY 2
order by 1 DESC;


D.

WITH range_quart AS
(SELECT country_name, region, percent_forest,
(CASE
WHEN percent_forest BETWEEN 75 AND 100 THEN 'Fourth'
WHEN percent_forest <= 75 AND percent_forest > 50 THEN 'Third'
WHEN percent_forest <= 50 AND percent_forest > 25 THEN 'Second'
ELSE 'First' END) AS range_quartiles
FROM forestation
WHERE year = 2016
AND region NOT LIKE 'World'
AND percent_forest IS NOT NULL)

SELECT country_name,region, percent_forest
FROM range_quart
WHERE range_quartiles = 'Fourth'
order by 3 DESC;
```

```sql
E.
SELECT COUNT(country_code) AS countries_higher
FROM forestation
WHERE year = 2016
AND percent_forest > (SELECT percent_forest
                      FROM forestation WHERE year = 2016 AND country_code = 'USA');
```