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 3.208 %.

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 1279999.99 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
Latin America & Caribbean	51.03	46.16
North America	35.65	36.04
Sub-Saharan Africa	30.67	28.79
Europe & Central Asia	37.28	38.04
Middle East & North Africa	1.78	2.07
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 527229.06 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.

United States and China 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 km^2
Indonesia	East Asia & Pacific	-282193 km^2
Myanmar	East Asia & Pacific	-107234 km^2
Nigeria	Sub-Saharan Africa	-106506 km^2
Tanzania	Sub-Saharan Africa	-102320 km^2

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
High: 75-100%	9
Medium: 50-75%	38
Low: 25-50%	73
Very low: 0-25%	85

The largest number of countries in 2016 were found in the **Very low: 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	98.26%
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?*
  - *Understand the status of world deforestation after 16 years.*
  - *Recognize that continents are experiencing serious and severe deforestation and especially in Africa*
- *Which countries should we focus on over others?*
  - *Countries located in Sub-Saharan Africa, 4 out of 5 highest countries are in this region.*

5. APPENDIX: SQL Queries Used

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

```
CREATE VIEW forestation AS
SELECT
    forest_area.*,
    regions.region,
    regions.income_group,
    land_area.total_area_sq_mi * 2.59 AS total_area_km,
    (forest_area.forest_area_sqkm / (land_area.total_area_sq_mi *
2.59))* 100 AS percentage_forest
FROM
    forest_area INNER JOIN land_area
    ON forest_area.country_code = land_area.country_code
    AND forest_area.year = land_area.year
    INNER JOIN regions
    ON forest_area.country_code = regions.country_code
ORDER BY
    country_code,year;
```

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.

```
SELECT forest_area_sqkm
FROM forestation
WHERE year = '1990'
AND country_name = 'World';
```

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 forest_area_sqkm
FROM forestation
WHERE year = '2016'
AND country_name = 'World';
```

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

```
SELECT (fa_1990.forest_area_sqkm - fa_2016.forest_area_sqkm) AS
forest_area_change
FROM forestation fa_1990
INNER JOIN forestation fa_2016
ON fa_1990.country_name = fa_2016.country_name
WHERE
    fa_1990.year = '1990'
    AND fa_1990.region = 'World'
    AND fa_2016.year = '2016'
    AND fa_2016.region = 'World';
```

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

```
WITH forest_area_1990 AS(
    SELECT forest_area_sqkm AS fa_km_1990
    FROM forestation
    WHERE year = '1990'
    AND country_name = 'World'
),forest_area_2016 AS(
    SELECT forest_area_sqkm AS fa_km_2016
    FROM forestation
    WHERE year = '2016'
    AND country_name = 'World'
)
SELECT ROUND((((forest_area_1990.fa_km_1990 -
forest_area_2016.fa_km_2016)
       /forest_area_1990.fa_km_1990)*100):: NUMERIC, 3
       ) as fa_change_percent
FROM forest_area_1990, forest_area_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?

```
WITH forest_area_1990 AS(
    SELECT forest_area_sqkm AS fa_km_1990
    FROM forestation
    WHERE year = '1990'
    AND country_name = 'World'
),forest_area_2016 AS(
    SELECT forest_area_sqkm AS fa_km_2016
    FROM forestation
    WHERE year = '2016'
    AND country_name = 'World'
)
SELECT country_name,total_area_km,
(forest_area_1990.fa_km_1990 - forest_area_2016.fa_km_2016)
AS forest_area_change_of_world
FROM forestation,forest_area_2016,forest_area_1990
WHERE year = '2016'
ORDER BY ABS(total_area_km - (forest_area_1990.fa_km_1990 -
forest_area_2016.fa_km_2016))
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 forest_2016 AS(
    SELECT region,
      SUM(forest_area_sqkm) AS total_forest_area_km,
      SUM(total_area_km) as total_land_area_km
    from forestation
    GROUP BY 1, year
    HAVING year = '2016')
SELECT forest_2016.*,
ROUND(
((forest_2016.total_forest_area_km/forest_2016.total_land_area_km
)*100)::NUMERIC, 2) AS forest_percent_2016
FROM forest_2016;
```

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 forest_1990 AS(
    SELECT region,
      SUM(forest_area_sqkm) AS total_forest_area_km,
      SUM(total_area_km) as total_land_area_km
    from forestation
    GROUP BY 1, year
    HAVING year = '1990')
SELECT forest_1990.*,
ROUND(((forest_1990.total_forest_area_km/forest_1990.total_land_a
rea_km)*100)::NUMERIC, 2)
    AS forest_percent_1990
    from forest_1990;
```

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

```
WITH per_1990 AS(
SELECT forest_1990.*,
ROUND(
((forest_1990.total_forest_area_km/forest_1990.total_land_area_km
)*100)::NUMERIC, 2)
    AS forest_percent_1990
    from(
        SELECT region,SUM(forest_area_sqkm) AS
total_forest_area_km, SUM(total_area_km) as total_land_area_km
        from forestation
        GROUP BY 1, year
        HAVING year = '1990'
    )as forest_1990
),per_2016 AS (
SELECT forest_2016.*,
ROUND(
((forest_2016.total_forest_area_km/forest_2016.total_land_area_km
)*100)::NUMERIC, 2)
```

```
    AS forest_percent_2016
    FROM (
        SELECT region,
           SUM(forest_area_sqkm) AS total_forest_area_km,
           SUM(total_area_km) as total_land_area_km
        FROM forestation
        GROUP BY 1, year
        HAVING year = '2016'
    )as forest_2016
)
SELECT per_1990.region, per_1990.forest_percent_1990,
per_2016.forest_percent_2016,
CASE
     WHEN per_1990.forest_percent_1990 >
     per_2016.forest_percent_2016
     THEN 'DECREASED'
ELSE 'INCREASED' END AS Status
FROM per_1990
    INNER JOIN per_2016 on per_1990.region = per_2016.region
    ORDER BY Status;
```

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 fa_km_1990 AS(
    SELECT *
    FROM forestation
    WHERE year = '1990'
),fa_km_2016 AS(
    SELECT *
    FROM forestation
    WHERE year = '2016'
)
SELECT fa_km_1990.country_name,
    fa_km_1990.region,
    fa_km_1990.forest_area_sqkm AS forest_area_km_1990,
    fa_km_2016.forest_area_sqkm AS forest_area_km_2016,
    (fa_km_2016.forest_area_sqkm - fa_km_1990.forest_area_sqkm)
AS change
FROM fa_km_1990
JOIN fa_km_2016
    on fa_km_1990.country_code = fa_km_2016.country_code
    WHERE fa_km_1990.region NOT LIKE 'World'
    ORDER BY change 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 fa_km_1990 AS(
    SELECT *
    FROM forestation
    WHERE year = '1990'
),fa_km_2016 AS(
    SELECT *
    FROM forestation
    WHERE year = '2016'
)
SELECT fa_km_1990.country_name,
    fa_km_1990.region,
    fa_km_1990.forest_area_sqkm AS forest_area_km_1990,
    fa_km_2016.forest_area_sqkm AS forest_area_km_2016,
    ROUND(-((1-
(fa_km_2016.forest_area_sqkm/fa_km_1990.forest_area_sqkm))*100)::
NUMERIC, 2
       ) AS fa_change_percent
FROM fa_km_1990
INNER JOIN fa_km_2016
    ON fa_km_1990.country_code = fa_km_2016.country_code
WHERE fa_km_1990.country_name NOT LIKE 'World'
    ORDER BY fa_change_percent
    LIMIT 5;
```

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

```
WITH percent_forest_quartiles AS(
    SELECT country_name,ROUND(percentage_forest::NUMERIC,2) AS
percentage_forest,
    CASE
     WHEN percentage_forest <= '25'
          THEN 'Very low: 0-25%'
     WHEN percentage_forest > '25' AND percentage_forest <= '50'
          THEN 'Low: 25-50%'
     WHEN percentage_forest > '50' AND percentage_forest <= '75'
          THEN 'Medium: 50-75%'
     ELSE 'High: 75-100%' END AS quartile_percent
    FROM forestation
    WHERE year = '2016' AND percentage_forest IS NOT NULL)
SELECT quartile_percent,COUNT(country_name) AS number_country
FROM percent_forest_quartiles GROUP BY 1
ORDER BY quartile_percent
DESC LIMIT 1;
```

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

```
WITH percent_forest_quartiles AS(
SELECT country_name,region,
ROUND(percentage_forest::NUMERIC,2) AS percentage_forest,
CASE WHEN percentage_forest <= '25'
       THEN 'Very low: 0-25%'
     WHEN percentage_forest > '25' AND percentage_forest <= '50'
       THEN 'Low: 25-50%'
     WHEN percentage_forest > '50' AND percentage_forest <= '75'
       THEN 'Medium: 50-75%'
     ELSE 'High: 75-100%' END AS quartile_percent
FROM forestation
WHERE year = '2016' AND percentage_forest IS NOT NULL
)
SELECT country_name,region, percentage_forest,quartile_percent
FROM percent_forest_quartiles
WHERE quartile_percent = 'High: 75-100%'
ORDER BY percentage_forest DESC;
```

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

```
SELECT COUNT(*) AS number_country_higher_US
FROM (
SELECT DISTINCT country_name
FROM forestation
WHERE percentage_forest >
       (SELECT percentage_forest
        FROM forestation
        WHERE country_name = 'United States' and year = '2016')
      ) AS count;
```