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 41,282,6945 km2 in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 39,958,246 km2, a loss of 1,324,449 km2, or 3.2 %.

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 1,279,999.99 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 46.16%, and the region with the lowest relative forestation was the Middle East and 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 and the Caribbean, with 51 %, and the region with the lowest relative forestation was 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
Europe and Central Asia	37.28	38.04
South Asia	16.51	17.51
East Asia & Pacific	25.78	26.36
Sub-Saharan Africa	30.67	28.79
North America	35.65	36.04
Middle East & North Africa	1.78	2.07

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**

- 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 km2. 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 km2, 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.

- 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	541,510 km2
Indonesia	East Asia & Pacific	282,193.98 km2

Myanmar	East Asia & Pacific	107,234 km2
Nigeria	Sub-Saharan Africa	106,506 km2
Tanzania	Sub-Saharan Africa	102,320 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	61.8
Uganda	Sub-Saharan Africa	59.27
Mauritania	Sub-Saharan Africa	46.75
Honduras	Latin America and 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 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.

- QUARTILES

Table 3.3: Count of Countries Grouped by Forestation Percent Quartiles, 2016:

Quartile	Number of Countries
75-100	9
50-75	38
25-50	72
0-25	85

The largest number of countries in 2016 were found in the last (85) 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	87.5
Guyana	Latin America & Caribbean	83.9
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?*

5. APPENDIX: SQL Queries Used

```
CREATE VIEW forestation AS
SELECT r.country_name,
    f.year,
    r.income_group,
    r.region,
    l.total_area_sq_mi * 2.59 AS total_area_sqkm,
    f.forest_area_sqkm,
    (f.forest_area_sqkm / (l.total_area_sq_mi * 2.59)) * 100 AS forest_percent
FROM forest_area AS f
JOIN land_area AS l ON f.country_code = l.country_code
AND f.year = l.year
JOIN regions AS r ON f.country_code = r.country_code
GROUP BY r.country_name,
     f.year,
     r.income_group,
     r.region,
     l.total_area_sq_mi,
     f.forest_area_sqkm;
```

Part 1 — Global Situation

a. What was the total forest area (in sq km) of the world in 1990?

```
SELECT ROUND(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?

```
SELECT ROUND(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

```
WITH world_forest_area AS (
    SELECT
        f1.forest_area_sqkm AS forest_area_sqkm_1990,
        f2.forest_area_sqkm AS forest_area_sqkm_2016
    FROM
        Forestation f1
    JOIN
        Forestation f2 ON f1.country_name = f2.country_name
    WHERE
        f1.YEAR = 1990 AND f2.YEAR = 2016 AND f1.country_name = 'World' AND
f2.country_name = 'World'
)
SELECT
    ROUND(forest_area_sqkm_1990 - forest_area_sqkm_2016) AS forest_area_change
```

```
FROM
    world_forest_area;
```

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

```
WITH world_forest_area AS (
    SELECT
        f1.forest_area_sqkm AS total_forest_area_1990,
        f2.forest_area_sqkm AS total_forest_area_2016
    FROM
        Forestation f1
    JOIN
        Forestation f2 ON f1.country_name = f2.country_name
    WHERE
        f1.YEAR = 1990 AND f2.YEAR = 2016 AND f1.country_name = 'World' AND
f2.country_name = 'World'
)
SELECT
    CAST(((total_forest_area_1990 - total_forest_area_2016) / total_forest_area_1990) * 100 AS
numeric(10,1)) AS percent_change_in_forest_area
FROM
    world_forest_area;
```

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_change AS
  (SELECT country_name,
          MAX(CASE
                WHEN YEAR = 2016 THEN total_area_sqkm
              END) AS total_area_2016,

     (SELECT MAX(forest_area_sqkm)
      FROM forestation
```

```
    WHERE YEAR = 1990
      AND country_name = 'World') AS forest_area_1990,

   (SELECT MAX(forest_area_sqkm)
    FROM forestation
    WHERE YEAR = 2016
      AND country_name = 'World') AS forest_area_2016
  FROM forestation
  GROUP BY country_name)
SELECT fc.country_name,
      fc.total_area_2016,
      fc.total_area_2016 - (fc.forest_area_1990 - fc.forest_area_2016) AS difference
FROM forest_change fc
ORDER BY ABS(fc.total_area_2016 - (fc.forest_area_1990 - fc.forest_area_2016)) ASC
LIMIT 1;
```

Part 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?*

```
SELECT CAST(SUM(forest_area_sqkm) / SUM(total_area_sqkm) * 100 AS NUMERIC)
AS percent_forest_world
FROM forestation
WHERE YEAR = 2016
  AND country_name = 'World';
```

- *Which region had the HIGHEST percent forest in 2016*

```
SELECT region,
      CAST(SUM(forest_area_sqkm) * 100 / SUM(total_area_sqkm) AS NUMERIC) AS
percent_forest
FROM forestation
WHERE YEAR = 2016
GROUP BY region
ORDER BY percent_forest DESC
LIMIT 1;
```

- *Which region had the LOWEST percent forest in 2016*

```
SELECT region,
      ROUND(CAST(SUM(forest_area_sqkm) * 100.0 / SUM(total_area_sqkm) AS
numeric), 2) AS percent_forest
FROM forestation
WHERE YEAR = 2016
GROUP BY region
ORDER BY percent_forest ASC
LIMIT 1;
```

b. *What was the percent forest of the entire world in 1990?*

```
SELECT forest_area_sqkm * 100/ total_area_sqkm AS total_forest_percent
FROM forestation
WHERE YEAR = 1990
 AND country_name = 'World' ;
```

- *Which region had the HIGHEST percent forest in 1990*

```
SELECT region,
      ROUND(SUM(forest_area_sqkm) * 100.0 / SUM(total_area_sqkm)) AS
percent_forest
FROM forestation
WHERE YEAR = 1990
  AND region NOT LIKE 'World'
GROUP BY region
ORDER BY percent_forest DESC
LIMIT 1;
```

- *Which region had the LOWEST percent forest in 1990*

```
SELECT region,
      ROUND(CAST(SUM(forest_area_sqkm) * 100.0 / SUM(total_area_sqkm) AS
numeric), 2) AS percent_forest
FROM forestation
WHERE YEAR = 1990
  AND region NOT LIKE 'World'
GROUP BY region
ORDER BY percent_forest ASC
LIMIT 1;
```

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

```
WITH forest_change AS

 (SELECT region,

       SUM(CASE

            WHEN YEAR = 1990 THEN forest_area_sqkm
```

```
            ELSE 0
        END) AS forest_sum_1990,
    SUM(CASE
            WHEN YEAR = 2016 THEN forest_area_sqkm
            ELSE 0
        END) AS forest_sum_2016
  FROM forestation
  WHERE region <> 'World'
  GROUP BY region)
SELECT region,
    ((forest_sum_2016 - forest_sum_1990) * 100.0 / NULLIF(forest_sum_1990, 0)) AS
percent_change
FROM forest_change
WHERE forest_sum_2016 < forest_sum_1990;
```

Part 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 forestation_data AS (
  SELECT
    f1.region,
    f1.country_name,
    SUM(CASE WHEN f1.year = 1990 THEN f1.forest_area_sqkm ELSE 0 END) AS
forest_sum_1990,
    SUM(CASE WHEN f2.year = 2016 THEN f2.forest_area_sqkm ELSE 0 END) AS
forest_sum_2016
  FROM
    forestation f1
  JOIN
    forestation f2 ON f1.region = f2.region AND f1.country_name = f2.country_name
  WHERE
    f1.region <> 'World' AND f2.region <> 'World'
  GROUP BY
    f1.region, f1.country_name
)
SELECT
```

```
    fd.region,
    fd.country_name,
    fd.forest_sum_1990 AS forest_1990,
    fd.forest_sum_2016 AS forest_2016,
    ROUND(CAST((fd.forest_sum_1990 - fd.forest_sum_2016) AS numeric), 2) AS
difference
FROM
    forestation_data fd
WHERE
    fd.forest_sum_2016 < fd.forest_sum_1990
ORDER BY
    difference 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 Forestation_1990 AS (
    SELECT
        country_name,
        (SUM(forest_area_sqkm) / SUM(total_area_sqkm * 2.59)) * 100 AS
percent_forestation_1
    FROM
        forestation
    WHERE
        year = 1990
    GROUP BY
        country_name
),
Forestation_2016 AS (
    SELECT
        country_name,
        (SUM(forest_area_sqkm) / SUM(total_area_sqkm * 2.59)) * 100 AS
percent_forestation_2
    FROM
        forestation
    WHERE
        year = 2016
    GROUP BY
        country_name
)
SELECT
    f.country_name,
    ROUND((((f.percent_forestation_1 - t.percent_forestation_2) / f.percent_forestation_1)
* 100)::numeric, 2) AS percent_change
```

```
FROM
   Forestation_1990 f
JOIN
   Forestation_2016 t ON f.country_name = t.country_name
WHERE
   f.percent_forestation_1 IS NOT NULL
   AND t.percent_forestation_2 IS NOT NULL
   AND f.country_name != 'World'
ORDER BY
   percent_change DESC
LIMIT
   5;
```

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

```
WITH forestation_quartiles AS
 (SELECT region,
        country_name,
        CASE
           WHEN forest_percent > 75 THEN 'Fourth'
           WHEN forest_percent <= 75
               AND forest_percent > 50 THEN 'Third'
           WHEN forest_percent <= 50
               AND forest_percent > 25 THEN 'Second'
           ELSE 'First'
        END AS quartiles
  FROM forestation
  WHERE YEAR = 2016
   AND region NOT LIKE 'World'
   AND forest_percent IS NOT NULL )
SELECT quartiles,
      COUNT(*) AS quartiles_groups
FROM forestation_quartiles
GROUP BY quartiles;
```

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

```
SELECT country_name,
      region,
      forest_percent
FROM forestation
WHERE YEAR = 2016
 AND forest_percent > 75
```

```
ORDER BY forest_percent DESC;
```

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

```
SELECT COUNT(country_name)
FROM forestation f1
WHERE YEAR = 2016
  AND forest_percent >
    (SELECT forest_percent
     FROM forestation f2
     WHERE f2.country_name = 'United States'
       AND f2.year = 2016 );
```