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 41282694.90 sqkm in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 39958245.9sqkm, a loss of 1324449, 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 1279999.9891sqkm).

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 & 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
Latin America & Caribbean	51.03	46.16
Europe & Central Asia	37.28	38.04
North America	35.65	36.04
World	32.42	31.38
Sub-Saharan Africa	30.67	28.79
East Asia & Pacific	25.78	26.36
South Asia	16.51	17.51
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 & 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 527229.06sqkm. 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 79200.00sqkm, 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.

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	541510.00
Indonesia	East Asia & Pacific	282193.98
Myanmar	East Asia & Pacific	107234.00
Nigeria	Sub-Saharan Africa	106506.00
Tanzania	Sub-Saharan Africa	102320.00

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	5168.00	75.45
Nigeria	106506.00	61.80
Uganda	28092.00	59.13
Mauritania	1940.00	46.75
Honduras	36640.00	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
First	85
Second	73
Third	38
Fourth	9

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

There were 94 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?*

Deforestation has been gradually happening. It's a real threat. The analysis showed a reduction of the global forest of 3.208% from 1990 to 2016. One of the most impacted regions is Sub-Sara Africa. For example, when we look at 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 Sub-Saharan Africa. This is followed by Latin America & Caribbean. These regions are developing countries and maybe they may not be aware of the danger of deforestation. Poverty is also prevalent in these regions, so we can say that the poorest regions of the world are experiencing higher deforestation than other regions.

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

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 needs to step up remedial efforts. Other countries that we should focus are Togo, Uganda, Brazil. The recommendation is that we should create awareness in this region educating them on the danger of deforestation. They need to understand that their way of life is having an ecological impact in their area.

5. APPENDIX: SQL Queries Used

Create a View called "forestation" by joining all three tables - forest_area, land_area, and regions

```
CREATE VIEW forestation AS
SELECT
   f.country_code,
   f.country_name,
   f.year,
   f.forest_area_sqkm,
   l.total_area_sq_mi,
   r.region,
   r.income_group,
   l.total_area_sq_mi * 2.59 AS total_area_sqkm,
   f.forest_area_sqkm / (l.total_area_sq_mi * 2.59) * 100 AS percent_of_land_area_that_is_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 f.country_code = r.country_code;
```

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.

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.”

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

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

```
WITH sub AS (
   SELECT year, country_name, region, forest_area_sqkm
   FROM forestation
   WHERE country_name = 'World' AND (year = 2016 OR year = 1990)
)
SELECT
   s1.region,
   s1.forest_area_sqkm AS forest_area_sqkm_2016,
   s2.forest_area_sqkm AS forest_area_sqkm_1990,
   s2.forest_area_sqkm - s1.forest_area_sqkm AS drop_in_forest_area,
   (s2.forest_area_sqkm - s1.forest_area_sqkm) * 100 / s2.forest_area_sqkm AS
percentage_drop_in_forest_area
FROM
   sub s1
JOIN
   sub s2
ON
   s1.region = s2.region AND s1.year > s2.year;
```

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 sub AS (
   SELECT year, country_name, region, forest_area_sqkm
   FROM forestation
   WHERE country_name = 'World' AND (year = 2016 OR year = 1990)
),
sub2 AS (
```

```
    SELECT
        s1.region,
        s1.forest_area_sqkm AS forest_area_sqkm_2016,
        s2.forest_area_sqkm AS forest_area_sqkm_1990,
        s2.forest_area_sqkm - s1.forest_area_sqkm AS drop_in_forest_area,
        (s2.forest_area_sqkm - s1.forest_area_sqkm) * 100 / s2.forest_area_sqkm AS
percentage_drop_in_forest_area
    FROM
        sub s1
    JOIN
        sub s2
    ON
        s1.region = s2.region AND s1.year > s2.year
)

SELECT
    year,
    country_name,
    total_area_sqkm,
    ABS(total_area_sqkm - (SELECT drop_in_forest_area FROM sub2)) AS absolute_diff
FROM
    forestation
WHERE
    year = 2016
ORDER BY
    absolute_diff
LIMIT 1;
```

Part 2 - Regional Outlook

Create a table that shows the Regions and their percent forest area (sum of forest area divided by the sum of land area) in 1990 and 2016. (Note that 1 sq mi = 2.59 sq km).
Based on the table you created:

```
SELECT
    region,
    SUM(forest_area_sqkm) AS forest_area_sqkm,
    SUM(total_area_sqkm) AS total_area_sqkm,
    ROUND(
        CAST((SUM(forest_area_sqkm) / SUM(total_area_sqkm)) * 100 AS NUMERIC),
        2
    ) AS percent_forest_area
```

```
FROM
   forestation
WHERE
   year = 2016
GROUP BY
   region
ORDER BY
   percent_forest_area DESC;
```

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
   region,
   SUM(forest_area_sqkm) AS forest_area_sqkm,
   SUM(total_area_sqkm) AS total_area_sqkm,
   ROUND(
      CAST((SUM(forest_area_sqkm) / SUM(total_area_sqkm)) * 100 AS NUMERIC),
      2
   ) AS percent_forest_area
FROM
   forestation
WHERE
   year = 2016 AND region = 'World'
GROUP BY
   region
```

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?

```
SELECT
   region,
   SUM(forest_area_sqkm) AS forest_area_sqkm,
   SUM(total_area_sqkm) AS total_area_sqkm,
   ROUND(
      CAST((SUM(forest_area_sqkm) / SUM(total_area_sqkm)) * 100 AS NUMERIC),
      2
   ) AS percent_forest_area
FROM
   forestation
WHERE
   year = 1990 AND region = 'World'
GROUP BY
   region
```

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

```
WITH sub AS (
   SELECT
      year,
      region,
      forest_area_sqkm,
      total_area_sqkm
   FROM
      forestation
   WHERE
      year = 2016 OR year = 1990
),
sub2 AS (
   SELECT
      s1.region,
      s1.forest_area_sqkm AS forest_area_sqkm_2016,
      s1.total_area_sqkm AS total_area_sqkm_2016,
      s2.forest_area_sqkm AS forest_area_sqkm_1990,
      s2.total_area_sqkm AS total_area_sqkm_1990
   FROM
      sub s1
   JOIN
      sub s2 ON s1.region = s2.region AND s1.year > s2.year
)

SELECT
   region,
   ROUND(
      CAST(SUM(forest_area_sqkm_1990) * 100 / SUM(total_area_sqkm_1990) AS NUMERIC),
      2
   ) AS perc_forest_area_1990,
   ROUND(
      CAST(SUM(forest_area_sqkm_2016) * 100 / SUM(total_area_sqkm_2016) AS NUMERIC),
      2
   ) AS perc_forest_area_2016
FROM
   sub2
GROUP BY
   region
ORDER BY
```

```
    perc_forest_area_1990 DESC;
```

Part 3 - Country-Level Detail

Success Stories

```
WITH sub AS (
    SELECT
        year,
        country_name,
        region,
        forest_area_sqkm
    FROM
        forestation
    WHERE
        (year = 2016 OR year = 1990) AND
        region != 'World' AND
        forest_area_sqkm IS NOT NULL
)

SELECT
    s1.country_name,
    s1.forest_area_sqkm AS forest_area_sqkm_2016,
    s2.forest_area_sqkm AS forest_area_sqkm_1990,
    ROUND(CAST((s1.forest_area_sqkm - s2.forest_area_sqkm) AS NUMERIC), 2) AS
increase_in_forest_area,
    ROUND(CAST(100 * ((s1.forest_area_sqkm - s2.forest_area_sqkm) / s2.forest_area_sqkm)
AS NUMERIC), 2) AS perc_change_in_forest_area
FROM
    sub s1
JOIN
    sub s2 ON s1.country_name = s2.country_name AND s1.year > s2.year
ORDER BY
    increase_in_forest_area DESC;


WITH sub AS (
    SELECT
        year,
        country_name,
        region,
        forest_area_sqkm
```

```
    FROM
        forestation
    WHERE
        (year = 2016 OR year = 1990) AND
        region != 'World' AND
        forest_area_sqkm IS NOT NULL
)

SELECT
    s1.country_name,
    s1.forest_area_sqkm AS forest_area_sqkm_2016,
    s2.forest_area_sqkm AS forest_area_sqkm_1990,
    ROUND(CAST((s1.forest_area_sqkm - s2.forest_area_sqkm) AS NUMERIC), 2) AS
increase_in_forest_area,
    ROUND(CAST(100 * ((s1.forest_area_sqkm - s2.forest_area_sqkm) / s2.forest_area_sqkm)
AS NUMERIC), 2) AS perc_change_in_forest_area
FROM
    sub s1
JOIN
    sub s2 ON s1.country_name = s2.country_name AND s1.year > s2.year
ORDER BY
    perc_change_in_forest_area DESC;
```

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 sub AS (
    SELECT
        year,
        country_name,
        region,
        forest_area_sqkm
    FROM
        forestation
    WHERE
        (year = 2016 OR year = 1990) AND
        region != 'World' AND
        forest_area_sqkm IS NOT NULL
)

SELECT
    s1.country_name,
    s1.region,
```

```
   ROUND(CAST((s2.forest_area_sqkm - s1.forest_area_sqkm) AS NUMERIC), 2) AS
change_in_forest_area
FROM
   sub s1
JOIN
   sub s2 ON s1.country_name = s2.country_name AND s1.year > s2.year
ORDER BY
   change_in_forest_area 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 sub AS (
   SELECT
      year,
      country_name,
      region,
      forest_area_sqkm
   FROM
      forestation
   WHERE
      (year = 2016 OR year = 1990) AND
      region != 'World' AND
      forest_area_sqkm IS NOT NULL
)

SELECT
   s1.country_name,
   s1.region,
   ROUND(CAST((s2.forest_area_sqkm - s1.forest_area_sqkm) AS NUMERIC), 2) AS
change_in_forest_area,
   ROUND(CAST(100 * ((s2.forest_area_sqkm - s1.forest_area_sqkm) / s2.forest_area_sqkm)
AS NUMERIC), 2) AS per_change_in_forest_area
FROM
   sub s1
JOIN
   sub s2 ON s1.country_name = s2.country_name AND s1.year > s2.year
ORDER BY
   per_change_in_forest_area DESC
LIMIT 5;
```

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

```
WITH sub AS (
    SELECT
        year,
        country_name,
        region,
        forest_area_sqkm * 100 / total_area_sqkm AS percent_forestation
    FROM
        forestation
    WHERE
        year = 2016
),
sub2 AS (
    SELECT
        country_name,
        CASE
            WHEN percent_forestation > 75 THEN 'Fourth'
            WHEN percent_forestation > 50 THEN 'Third'
            WHEN percent_forestation > 25 THEN 'Second'
            ELSE 'First'
        END AS quartile_category
    FROM
        sub
    WHERE
        percent_forestation IS NOT NULL
)

SELECT
    DISTINCT quartile_category,
    COUNT(country_name) OVER(PARTITION BY quartile_category) AS number_of_countries
FROM
    sub2
ORDER BY
    number_of_countries DESC;
```

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

```
WITH sub AS (
    SELECT
        year,
        country_name,
```

```
        region,
        ROUND(CAST(forest_area_sqkm * 100 / total_area_sqkm AS NUMERIC), 2) AS
percent_forestation
    FROM
        forestation
    WHERE
        year = 2016
),
sub2 AS (
    SELECT
        country_name,
        region,
        percent_forestation,
        CASE
            WHEN percent_forestation > 75 THEN 'Fourth'
            WHEN percent_forestation > 50 THEN 'Third'
            WHEN percent_forestation > 25 THEN 'Second'
            ELSE 'First'
        END AS quartile_category
    FROM
        sub
    WHERE
        percent_forestation IS NOT NULL
)

SELECT
    country_name,
    region,
    percent_forestation
FROM
    sub2
WHERE
    quartile_category = 'Fourth'
ORDER BY
    percent_forestation DESC;
```

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

```
WITH sub AS (
    SELECT
        year,
        country_name,
        region,
```

```
    ROUND(CAST(forest_area_sqkm * 100 / total_area_sqkm AS NUMERIC), 2) AS
percent_forestation
  FROM
    forestation
  WHERE
    year = 2016
),
sub2 AS (
  SELECT
    country_name,
    percent_forestation,
    CASE
      WHEN percent_forestation > 75 THEN 'Fourth'
      WHEN percent_forestation > 50 THEN 'Third'
      WHEN percent_forestation > 25 THEN 'Second'
      ELSE 'First'
    END AS quartile_category
  FROM
    sub
  WHERE
    percent_forestation IS NOT NULL
)

SELECT
  COUNT(*)
FROM
  sub2
WHERE
  percent_forestation > (
    SELECT
      percent_forestation
    FROM
      sub2
    WHERE
      country_name = 'United States'
  );
```