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 41,282,694.9 square kilometers in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 39958245.9 square kilometers, a loss of 1,324,449 square kilometers, 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 1,279,999.9891 square kilometers).

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 46.14%, 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 51.08 %, 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	51.08	46.14
Middle East & North Africa	1.78	2.07
North America	35.66	36.02

South Asia	16.53	17.50
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 American & Pacific (dropped from 51.08% to 46.14 %) and Sub-Suharan 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 527,229.062 square kilometers. 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 square kilometers, much lower than the figure for China.

China and the 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	541,510
Indonesia	East Asia & Pacific	282,193.9844

Myanmar	East Asia & Pacific	107,234.0039
Nigeria	Sub-Saharan Africa	106,506.00098
Tanzania	Sub-Saharan Africa	102,320

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 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
1	84
2	74
3	38
4	9

The largest number of countries in 2016 were found in the 1st 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	95.39
Palau	East Asia & Pacific	90.98
Gabon	Sub-Saharan Africa	90.04
Seychelles	Sub-Saharan Africa	88.41
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

5. RECOMMENDATIONS

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

- *What have you learned from the World Bank data?*

The World Bank data allowed a holistic view on the world's deforestation situation through global, regional and country level. Forest areas, forestation percentage, countries raising concerns over deforestation, etc. are thoroughly investigated by using SQL queries to explore and analyze the World Bank data.

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

The report shows that top concern-raising countries are Togo, Nigeria, Uganda, Mauritania and Honduras. These countries are mainly in Sub-Saharan Africa region; hence, further investigation in this region should be made with an aim to improve the deforestation situation.

6. APPENDIX: SQL QUERIES USED

1. Forestation View

```
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,
(
      (f.forest_area_sqkm / (l.total_area_sq_mi * 2.59)) * 100
   ) AS 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;
```

2. 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
      r.region region,
      SUM(f.forest_area_sqkm) forest_area
   FROM
      forest_area f
      JOIN regions r ON f.country_code = r.country_code
   WHERE
```

```
    r.region = 'World'
    AND f.year = '1990'
GROUP BY
    1
```

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
    r.region region,
    SUM(f.forest_area_sqkm) forest_area
FROM
    forest_area f
    JOIN regions r ON f.country_code = r.country_code
WHERE
    r.region = 'World'
    AND f.year = '2016'
GROUP BY
    1
```

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

```
WITH area_forest AS (
SELECT
    f.year,
    r.region region,
    SUM(f.forest_area_sqkm) forest_area
FROM
    forest_area f
    JOIN regions r ON f.country_code = r.country_code
WHERE
    f.year = '1990'
    OR f.year = '2016'
GROUP BY
    1,
    2
ORDER BY
```

```
    2,
    1
)
SELECT
  *,
  LEAD(forest_area) OVER (
    ORDER BY
      forest_area
  ) AS lead,
  LEAD(forest_area) OVER (
    ORDER BY
      forest_area
  ) - forest_area AS area_diff
FROM
  area_forest
WHERE
  region = 'World'
```

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

```
WITH area_forest AS (
  SELECT
    f.year,
    r.region region,
    SUM(f.forest_area_sqkm) forest_area
  FROM
    forest_area f
    JOIN regions r ON f.country_code = r.country_code
  WHERE
    f.year = '1990'
    OR f.year = '2016'
  GROUP BY
    1,
    2
  ORDER BY
    2,
    1
```

```
)
SELECT
  *,
  LEAD(forest_area) OVER (
    ORDER BY
      forest_area
  ) AS lead,
  (
    forest_area - LEAD(forest_area) OVER (
      ORDER BY
        forest_area
    )
  ) * 100 / LEAD(forest_area) OVER (
    ORDER BY
      forest_area
  ) AS percent_diff
FROM
  area_forest
WHERE
  region = 'World'
```

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?

```
SELECT
  country_name,
  SUM(total_area_sq_mi) * 2.59
FROM
  land_area
WHERE
  year = '2016'
GROUP BY
  1
HAVING
  1100000 <= SUM(total_area_sq_mi) * 2.59
  AND SUM(total_area_sq_mi) * 2.59 <= 1350000
ORDER BY
  2 DESC
```

3. 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
    r.region,
    f.year percent_year,
    SUM(f.forest_area_sqkm) forest_area_sqkm,
    SUM(l.total_area_sq_mi) total_area_sq_mi,
    CAST(
        (SUM(f.forest_area_sqkm) * 100) / (SUM(l.total_area_sq_mi) * 2.59) AS DECIMAL(10, 2)
    ) AS percent_forest_area
FROM
    regions r
    JOIN forest_area f ON r.country_code = f.country_code
    JOIN land_area l ON f.country_code = l.country_code
WHERE
    f.year = '2016'
GROUP BY
    1,
    2
ORDER BY
    5 DESC
```

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_region AS (
    SELECT
        r.region,
        f.year percent_year,
        SUM(f.forest_area_sqkm) forest_area_sqkm,
        SUM(l.total_area_sq_mi) total_area_sq_mi,
        CAST(
            (SUM(f.forest_area_sqkm) * 100) / (SUM(l.total_area_sq_mi) * 2.59) AS DECIMAL(10, 2)
        ) AS percent_forest_area,
        ROW_NUMBER() OVER(
```

```
        ORDER BY
          CAST(
            (SUM(f.forest_area_sqkm) * 100) / (SUM(l.total_area_sq_mi) * 2.59) AS DECIMAL(10, 2)
          )
      ) AS rank
    FROM
      regions r
      JOIN forest_area f ON r.country_code = f.country_code
      JOIN land_area l ON f.country_code = l.country_code
    WHERE
      f.year = '1990'
    GROUP BY
      1,
      2
)
SELECT
  *
FROM
  forest_region
WHERE
  rank = (
    SELECT
      MAX(rank)
    FROM
      forest_region
  )
  OR rank = (
    SELECT
      MIN(rank)
    FROM
      forest_region
  )
```

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

```
WITH forest_90 AS (
  SELECT
```

```
    r.region,
    f.year percent_year,
    SUM(f.forest_area_sqkm) forest_area_sqkm,
    SUM(l.total_area_sq_mi) total_area_sq_mi,
    CAST(
      (SUM(f.forest_area_sqkm) * 100) / (SUM(l.total_area_sq_mi) * 2.59) AS DECIMAL(10, 2)
    ) AS percent_forest_90
  FROM
    regions r
    JOIN forest_area f ON r.country_code = f.country_code
    JOIN land_area l ON f.country_code = l.country_code
  WHERE
    f.year = '1990'
  GROUP BY
    1,
    2
),
forest_16 AS (
  SELECT
    r.region,
    f.year percent_year,
    SUM(f.forest_area_sqkm) forest_area_sqkm,
    SUM(l.total_area_sq_mi) total_area_sq_mi,
    CAST(
      (SUM(f.forest_area_sqkm) * 100) / (SUM(l.total_area_sq_mi) * 2.59) AS DECIMAL(10, 2)
    ) AS percent_forest_16
  FROM
    regions r
    JOIN forest_area f ON r.country_code = f.country_code
    JOIN land_area l ON f.country_code = l.country_code
  WHERE
    f.year = '2016'
  GROUP BY
    1,
    2
)
SELECT
```

```
    forest_90.region,
    forest_90.percent_forest_90,
    forest_16.percent_forest_16
FROM
    forest_90
    JOIN forest_16 ON forest_90.region = forest_16.region
GROUP BY
    1,
    2,
    3
HAVING
    forest_90.percent_forest_90 - forest_16.percent_forest_16 > 0
```

4. 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 forest_area_90 AS (
    SELECT
        country_code,
        country_name,
        SUM(forest_area_sqkm) area_90
    FROM
        forest_area
    WHERE
        year = '1990'
    GROUP BY
        1,
        2
),
forest_area_16 AS (
    SELECT
        country_code,
        country_name,
        SUM(forest_area_sqkm) area_16
    FROM
        forest_area
```

```
  WHERE
    year = '2016'
  GROUP BY
    1,
    2
)
SELECT
  regions.region,
  forest_area_90.country_code,
  forest_area_16.country_name,
  forest_area_90.area_90 - forest_area_16.area_16 forest_diff
FROM
  regions
  JOIN forest_area_90 ON regions.country_code = forest_area_90.country_code
  JOIN forest_area_16 ON forest_area_90.country_code = forest_area_16.country_code
WHERE
  forest_area_90.area_90 - forest_area_16.area_16 IS NOT NULL
ORDER BY
  4 DESC
LIMIT
  6
```

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 forest_area_90 AS (
  SELECT
    country_code,
    country_name,
    SUM(forest_area_sqkm) area_90
  FROM
    forest_area
  WHERE
    year = '1990'
  GROUP BY
    1,
    2
),
```

```
forest_area_16 AS (
    SELECT
        country_code,
        country_name,
        SUM(forest_area_sqkm) area_16
    FROM
        forest_area
    WHERE
        year = '2016'
    GROUP BY
        1,
        2
)
SELECT
    regions.region,
    forest_area_90.country_code,
    forest_area_16.country_name,
    CAST(
        (forest_area_16.area_16 - forest_area_90.area_90) * 100 / forest_area_90.area_90 AS DECIMAL(10, 2)
    ) AS percent_change
FROM
    regions
    JOIN forest_area_90 ON regions.country_code = forest_area_90.country_code
    JOIN forest_area_16 ON forest_area_90.country_code = forest_area_16.country_code
ORDER BY
    4
LIMIT
    5
```

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

```
WITH forestation_percent AS (
    SELECT
        f.country_code country_code,
        f.country_name country_name,
        CAST(
            (SUM(f.forest_area_sqkm) * 100) / (SUM(l.total_area_sq_mi) * 2.59) AS DECIMAL(10, 2)
```

```
      ) AS forest_percent
    FROM
      forest_area f
      JOIN land_area l ON f.country_code = l.country_code
    WHERE
      f.year = '2016'
    GROUP BY
      1,
      2
)
SELECT
  standard_quartile,
  COUNT(*)
FROM
  (
    SELECT
      country_code,
      country_name,
      forest_percent,
      CASE
        WHEN forest_percent < 25 THEN '1st'
        WHEN forest_percent >= 25
        AND forest_percent < 50 THEN '2nd'
        WHEN forest_percent >= 50
        AND forest_percent < 75 THEN '3rd'
        WHEN forest_percent >= 75 THEN '4th'
        ELSE '1st'
      END AS standard_quartile
    FROM
      forestation_percent
    WHERE
      forest_percent IS NOT NULL
  ) quartiles
GROUP BY
  1
ORDER BY
  2 DESC
```

```
LIMIT
  1
```

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

```
WITH forestation_percent AS (
  SELECT
    f.country_code country_code,
    f.country_name country_name,
    CAST(
      (SUM(f.forest_area_sqkm) * 100) / (SUM(l.total_area_sq_mi) * 2.59) AS DECIMAL(10, 2)
    ) AS forest_percent
  FROM
    forest_area f
    JOIN land_area l ON f.country_code = l.country_code
  WHERE
    f.year = '2016'
  GROUP BY
    1,
    2
)
SELECT
  *
FROM
  (
    SELECT
      country_code,
      country_name,
      forest_percent,
      CASE
        WHEN forest_percent < 25 THEN '1st'
        WHEN forest_percent >= 25
        AND forest_percent < 50 THEN '2nd'
        WHEN forest_percent >= 50
        AND forest_percent < 75 THEN '3rd'
        WHEN forest_percent >= 75 THEN '4th'
        ELSE '1st'
```

```
        END AS standard_quartile
    FROM
        forestation_percent
    WHERE
        forest_percent IS NOT NULL
  ) quartiles
WHERE
  standard_quartile = '4th'
ORDER BY
  forest_percent DESC
```

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

```
WITH forest_16 AS (
  SELECT
    f.country_code,
    f.country_name,
    CAST(
      (SUM(f.forest_area_sqkm) * 100) / (SUM(l.total_area_sq_mi) * 2.59) AS DECIMAL(10, 2)
    ) AS percent_forest_16
  FROM
    forest_area f
    JOIN land_area l ON f.country_code = l.country_code
  WHERE
    f.year = '2016'
  GROUP BY
    1,
    2
)
SELECT
  COUNT(*)
FROM
  forest_16
WHERE
  percent_forest_16 > (
    SELECT
      percent_forest_16
```

```
    FROM
        forest_16
    WHERE
        country_code = 'USA'
)
```