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.**CREATING THE VIEW**

2. **GLOBAL SITUATION**

According to the World Bank, the total forest area of the world was 41,282,294.9 sqkm in 1990.

As of 2016, the most recent year for which data was available, that number had fallen to 39,958,245.9 sqkm , a loss of 1,324,449 sqkm, 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 1279999.9891 sqkm).

3. **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 3.1: 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 (dropped from 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%.

4. **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 527229.062 sqkm. 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 sqkm, 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 4.1: Top 5 Amount Decrease in Forest Area by Country, 1990 & 2016:

Country	Region	Absolute Forest Area Change
Brazil	Latin America and Caribbean	541,510.00 sqkm
Indonesia	East Asia and Pacific	282193.98 sqkm
Myanmar	East Asia and Pacific	107234.00 sqkm
Nigeria	Sub-Saharan Africa	106506.00 sqkm
Tanzania	Sub-Saharan Africa	102320.00 sqkm

The second way to consider which countries are of concern is to analyze the data by percent decrease.

Table 4.2: 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	59.27%
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 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 4.3: Count of Countries Grouped by Forestation Percent Quartiles, 2016:

Quartile	Number of Countries
0-25	85
25-50	73
50-75	38
75-100	9

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

There were 85 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 4.4: Top Quartile Countries, 2016:

Country	Region	Pct Designated as Forest
Suriname	Latin America and Caribbean	98.26
Micronesia, Fed. Sts.	East Asia and Pacific	91.86
Gabon	Sub-Saharan Africa	90.04
Seychelles	Sub-Saharan Africa	88.41
Palau	East Asia and Pacific	87.61
American Samoa	East Asia and Pacific	87.50
Guyana	Latin America and Caribbean	83.90

Lao PDR	East Asia and Pacific	82.11
Solomon Islands	East Asia and Pacific	77.86

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

country_name	quartiles	percent_forest	count
Suriname	75-100	98.26	9
Micronesia, Fed. Sts.	75-100	91.86	9
Gabon	75-100	90.04	9
Seychelles	75-100	88.41	9
Palau	75-100	87.61	9
American Samoa	75-100	87.50	9
Guyana	75-100	83.90	9
Lao PDR	75-100	82.11	9
Solomon Islands	75-100	77.86	9

4e. How many countries had a percent forestation higher than the United States in 2016?

- There were 94 Countries that had a percent forestation higher than the United States in 2016

5. 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?*

Based on the dataset and analysis, it is apparent the only two Regions had a decrease in forest area, Latin America and the Caribbean & Sub-Saharan Africa while the rest of the World saw an increase. However, I would like to point your attention to Table 4.1: Top 5 Amount Decrease in Forest Area by Country, 1990 & 2016. These countries saw a significant decrease in forest area and we should pay close attention to them in the coming years. I would recommend a plantation plan as well as a deeper look at each country's forest planning and best practices, if any.

Appendix

1)

```
CREATE VIEW forestation
AS
  SELECT r.country_code,
         r.country_name,
         f.year,
         r.income_group,
         r.region,
         l.total_area_sq_mi,
         f.forest_area_sqkm,
         ( ( Sum(forest_area_sqkm) / Sum(total_area_sq_mi * 2.59) ) * 100 )
         forest_percentage
  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
  GROUP  BY 1,
            2,
            3,
            4,
            5,
            6,
            7
```

2.1)

```
SELECT country_name,
       year,
       Sum(forest_area_sqkm) total_forest_area
FROM   forestation
WHERE  year = 1990
       AND country_name = 'World'
GROUP  BY 1,
                  2
```

RESULT:

World 1990 41,282,694.9 sqkm

2.2a

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

RESULT:
World 2016 39,958,245.9 sqkm

2.2b)

```
SELECT ( (SELECT Sum(forest_area_sqkm) total_forest_area
         FROM   forestation
         WHERE  year = 1990
                AND country_name = 'World') - (SELECT
                Sum(forest_area_sqkm) total_forest_area
                                       FROM   forestation
                                       WHERE  year = 2016
                                              AND country_name =
                                                  'World') ) AS
       forest_area_lost
FROM   forestation
LIMIT  1
```

RESULT:
difference = 1,324,449sqkm

2.2c)

```
WITH percent_change AS
(
       SELECT (((
              (
                     SELECT Sum(forest_area_sqkm) total_forest_area
                     FROM   forestation
                     WHERE  year = 1990
                     AND    country_name = 'World') -
              (
                     SELECT Sum(forest_area_sqkm) total_forest_area
                     FROM   forestation
                     WHERE  year = 2016
                     AND    country_name = 'World')) / (
              (
                     SELECT Sum(forest_area_sqkm) total_forest_area
                     FROM   forestation
                     WHERE  year = 1990
                     AND    country_name = 'World'))) *100) AS percent_diff
       FROM   forestation LIMIT 1)
SELECT Concat(Round(p.percent_diff::numeric,2), '%') AS percent_diff
FROM   percent_change p
```

RESULT

percent_diff = 3.21%

2.3a)

```
WITH country_closest AS
(
        SELECT   country_name,
                 Sum(total_area_sq_mi*2.59) total_land_area
        FROM     forestation
        WHERE    year = 2016
        GROUP BY 1
        ORDER BY 2 )
SELECT   c.country_name,
        c.total_land_area
FROM     country_closest c
WHERE    c.total_land_area < '1314449'
ORDER BY 2 DESC limit 1
```

RESULT:
Peru total_land_area= 1279999.9891 sqkm

3a & 3.1)

```
WITH year_1990
    AS (SELECT region,
               SUM(total_area_sq_mi)
               AS
                 total_area_sq_mi,
               SUM(forest_area_sqkm)
               AS
                 forest_area_sqkm,
               ( ( SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59) ) * 100
               ) AS
               forest_percentage_1990
        FROM   forestation
        WHERE  year = '1990'
        GROUP  BY 1
        ORDER  BY 1),
    year_2016
    AS (SELECT region,
               SUM(total_area_sq_mi)
               AS
                 total_area_sq_mi,
               SUM(forest_area_sqkm)
               AS
                 forest_area_sqkm,
               ( ( SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59) ) * 100
               ) AS
               forest_percentage_2016
        FROM   forestation
        WHERE  year = '2016'
        GROUP  BY 1
```

```
      ORDER  BY 1)
SELECT s.region,
     Round(s.forest_percentage_1990 :: NUMERIC, 2) AS forest_percentage_1990,
     Round(e.forest_percentage_2016 :: NUMERIC, 2) AS forest_percentage_2016
FROM   year_1990 s
     join year_2016 e
       ON e.region = s.region
ORDER  BY 2 DESC

4a)
WITH t1
   AS (SELECT country_name,
            Sum(forest_area_sqkm) AS forest_area_sqkm_1990
       FROM   forestation
       WHERE  year = '1990'
            AND forest_area_sqkm IS NOT NULL
       GROUP  BY 1
       ORDER  BY 2 DESC),
   t2
   AS (SELECT country_name,
            Sum(forest_area_sqkm) AS forest_area_sqkm_2016
       FROM   forestation
       WHERE  year = '2016'
            AND forest_area_sqkm IS NOT NULL
       GROUP  BY 1
       ORDER  BY 2 DESC)
SELECT t2.country_name,
     t1.forest_area_sqkm_1990,
     t2.forest_area_sqkm_2016,
     ( t2.forest_area_sqkm_2016 - t1.forest_area_sqkm_1990 ) AS
     gained_forest_area
FROM   t2
     JOIN t1
       ON t2.country_name = t1.country_name
WHERE  t1.forest_area_sqkm_1990 < t2.forest_area_sqkm_2016
     AND t2.country_name != 'World'
ORDER  BY 4 DESC

4b)
WITH t1
   AS (SELECT country_name,
            ( SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59) ) * 100
            percent_forestation_1990
       FROM   forestation
       WHERE  year = 1990
       GROUP  BY country_name,
                 forest_area_sqkm),
   t2
   AS (SELECT country_name,
            ( SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59) ) * 100
```

```
            AS
            percent_forestation_2016
        FROM   forestation
        WHERE  year = 2016
        GROUP  BY country_name,
                  forest_area_sqkm)
SELECT t1.country_name,
      Round(( ( ( t2.percent_forestation_2016 - t1.percent_forestation_1990
) / ( t1.percent_forestation_1990 ) ) * 100 ) :: NUMERIC, 2)
percent_change
FROM   t1
      join t2
        ON t1.country_name = t2.country_name
WHERE  t1.percent_forestation_1990 IS NOT NULL
      AND t2.percent_forestation_2016 IS NOT NULL
      AND t1.country_name != 'World'
ORDER  BY percent_change DESC

4.1)
WITH t1 AS
(
        SELECT   country_name,
                 region,
                 Sum(forest_area_sqkm) AS forest_area_sqkm_1990
        FROM     forestation
        WHERE    year = '1990'
        AND      forest_area_sqkm IS NOT NULL
        GROUP BY 1
        ORDER BY 2 DESC), t2 AS
(
        SELECT   country_name,
                 Sum(forest_area_sqkm) AS forest_area_sqkm_2016
        FROM     forestation
        WHERE    year = '2016'
        AND      forest_area_sqkm IS NOT NULL
        GROUP BY 1
        ORDER BY 2 DESC)
SELECT   t2.country_name,
        t1.forest_area_sqkm_1990,
        t2.forest_area_sqkm_2016,
        (t1.forest_area_sqkm_1990 - t2.forest_area_sqkm_2016) AS lost_forest_area
FROM     t2
JOIN     t1
ON       t2.country_name = t1.country_name
WHERE    t1.forest_area_sqkm_1990 > t2.forest_area_sqkm_2016
AND      t2.country_name != 'World'
ORDER BY 4 DESC limit 5

4.2)
WITH t1
```

```
    AS (SELECT country_name,
               ( SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59) ) * 100
               percent_forestation_1990
        FROM   forestation
        WHERE  year = 1990
        GROUP  BY country_name,
                  forest_area_sqkm),
    t2
    AS (SELECT country_name,
               ( SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59) ) * 100
               AS
               percent_forestation_2016
        FROM   forestation
        WHERE  year = 2016
        GROUP  BY country_name,
                  forest_area_sqkm)
SELECT t1.country_name,
       Round(
( ( ( t1.percent_forestation_1990 -
t2.percent_forestation_2016 ) / ( t1.percent_forestation_1990 ) ) * 100 ) :: NUMERIC, 2) percent_change
FROM   t1
       join t2
         ON t1.country_name = t2.country_name
WHERE  t1.percent_forestation_1990 IS NOT NULL
       AND t2.percent_forestation_2016 IS NOT NULL
       AND t1.country_name != 'World'
ORDER  BY percent_change DESC

4.3)
WITH t1
    AS (SELECT country_name,
               year,
               ( Sum(forest_area_sqkm) / Sum(total_area_sq_mi * 2.59) ) * 100
               AS
               percent_forestation
        FROM   forestation
        WHERE  year = 2016
        GROUP  BY country_name,
                  year,
                  forest_area_sqkm)
SELECT DISTINCT( quartiles ),
               Count(country_name)
                OVER(
                  partition BY quartiles)
FROM   (SELECT country_name,
               CASE
                 WHEN percent_forestation < 25 THEN '0-25'
                 WHEN percent_forestation >= 25
                      AND percent_forestation < 50 THEN '25-50'
                 WHEN percent_forestation >= 50
```

```
                AND percent_forestation < 75 THEN '50-75'
            ELSE '75-100'
          END AS quartiles
    FROM   t1
    WHERE  percent_forestation IS NOT NULL
          AND year = 2016
          AND country_name != 'World') sub


4.4)
WITH t1
   AS (SELECT country_name,
          year,
          ( SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59) ) * 100
          AS
            percent_forest
      FROM   forestation
      WHERE  year = '2016'
      GROUP  BY 1,
            2,
            forest_area_sqkm),
   t2
   AS (SELECT t1.country_name,
          t1.year,
          t1.percent_forest,
          CASE
            WHEN t1.percent_forest < 25 THEN '0-25'
            WHEN t1.percent_forest >= 25
                AND t1.percent_forest < 50 THEN '25-50'
            WHEN t1.percent_forest >= 50
                AND t1.percent_forest < 75 THEN '50-75'
            ELSE '75-100'
          END AS quartiles
      FROM   t1
      WHERE  t1.percent_forest IS NOT NULL),
   t3
   AS (SELECT t2.country_name,
          t2.quartiles,
          t2.percent_forest,
          Count(t2.country_name)
            over (
              PARTITION BY t2.quartiles)
      FROM   t2
      GROUP  BY 1,
            2,
            3)
SELECT t3.country_name,
     t3.quartiles,
     Round(( t3.percent_forest :: NUMERIC ), 2),
     t3.count
```

```
FROM   t3
WHERE  t3.quartiles = '75-100'
ORDER  BY 3 DESC
```

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

```
WITH t1
    AS (SELECT country_name,
               year,
               ( SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59) ) * 100
               AS
                 percent_forest
        FROM   forestation
        WHERE  year = '2016'
        GROUP  BY 1,
                  2,
                  forest_area_sqkm),
    t2
    AS (SELECT t1.country_name,
               t1.year,
               t1.percent_forest,
               CASE
                 WHEN t1.percent_forest < 25 THEN '0-25'
                 WHEN t1.percent_forest >= 25
                      AND t1.percent_forest < 50 THEN '25-50'
                 WHEN t1.percent_forest >= 50
                      AND t1.percent_forest < 75 THEN '50-75'
                 ELSE '75-100'
               END AS quartiles
        FROM   t1
        WHERE  t1.percent_forest IS NOT NULL),
    t3
    AS (SELECT t2.country_name,
               t2.quartiles,
               t2.percent_forest,
               Count(t2.country_name)
                 over (
                   PARTITION BY t2.quartiles)
        FROM   t2
        GROUP  BY 1,
                  2,
                  3)
SELECT t3.country_name,
       t3.quartiles,
       Round(( t3.percent_forest :: NUMERIC ), 2),
       t3.count
FROM   t3
WHERE  t3.quartiles = '75-100'
ORDER  BY 3 DESC
```

4e)
How many countries had a percent forestation higher than the United States in 2016?

```
WITH t1
    AS (SELECT country_name,
               ( ( Sum(forest_area_sqkm) / Sum(total_area_sq_mi * 2.59) ) * 100
               ) AS
               forest_percentage_2016
        FROM   forestation
        WHERE  year = '2016'
        GROUP  BY 1
        ORDER  BY 2 DESC),
    final
    AS (SELECT t1.country_name,
               t1.forest_percentage_2016,
               (SELECT t1.forest_percentage_2016 AS us_forest_percentage_2016
                FROM   t1
                WHERE  t1.country_name = 'United States')
        FROM   t1),
    final1
    AS (SELECT f.*,
               Count(*)
        FROM   final f
        GROUP  BY 1,
                  2,
                  3
        HAVING( f.forest_percentage_2016 ) > f.us_forest_percentage_2016
        ORDER  BY 2 DESC)
SELECT Concat(Sum(final1.count),
' Countries in 2016 had a percent forestation higher than the United States '
)
FROM   final1
```

RESOURCES.

Udacity
w3Schools
Youtube
Masterschool - SQL sessions