Report for ForestQuery into Global Deforestation, 1990 to 2016

SQL by Renee Raven

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 sq km** in 1990. As of 2016, the most recent year for which data was available, that number had fallen to **39,958,245.9 sq km**, a loss of **1,324,449 sq km**, 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 **1,280,000 sq km**).

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
World	32.42%	31.38%
Latin America & Caribbean	51.03%	46.16%
Europe & Central Asia	37.28%	38.04%
North America	35.65%	36.04%
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 **527229.06 sq km**. 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 **79200.00 sq km**, 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	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	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
1st <=25%	85
2nd >25% & <=50%	72
3rd >50% & <=75%	38
4th >75%	9

The largest number of countries in 2016 were found in the **1st (<=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

94 countries had a percent forestation higher than the United States in 2016.

4. SQL QUERIES

1. Create a View called “forestation” by joining all three tables - forest_area, land_area and regions in the workspace.
2. The forest_area and land_area tables join on both country_code AND year.
3. The regions table joins these based on only country_code.
4. In the ‘forestation’ View, include the following:
   - All of the columns of the origin tables
   - A new column that provides the percent of the land area that is designated as forest.
5. Keep in mind that the column forest_area_sqkm in the forest_area table and the land_area_sqmi in the land_area table are in different units (square kilometers and square miles, respectively), so an adjustment will need to be made in the calculation you write (1 sq mi = 2.59 sq km).

```
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
percent_designated_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;
```

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.

```sql
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.”

```sql
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?

```sql
WITH yr1990 AS
(
    SELECT forest_area_sqkm AS area1990, year, country_code
    FROM forestation
    WHERE year = '1990' AND country_name = 'World'
),
yr2016 AS
(
    SELECT forest_area_sqkm AS area2016, year, country_code
    FROM forestation
    WHERE year = '2016' AND country_name = 'World'
)
SELECT (area1990 - area2016) AS area_change
FROM yr1990
JOIN yr2016
ON yr1990.country_code = yr2016.country_code;
```

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

```sql
  WITH yr1990 AS
  (
      SELECT forest_area_sqkm AS area1990, year, country_code
      FROM forestation
      WHERE year = '1990' AND country_name = 'World'
  ),
  yr2016 AS
  (
```

```
      SELECT forest_area_sqkm AS area2016, year, country_code
      FROM forestation
      WHERE year = '2016' AND country_name = 'World'
  )
  SELECT ROUND(CAST(((area1990 - area2016)/area1990)*100 AS NUMERIC), 2) AS
  forest_change_percent
  FROM yr1990
  JOIN yr2016
  ON yr1990.country_code = yr2016.country_code;
```

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,
      ROUND(CAST((total_area_sq_mi * 2.59) AS numeric)) AS sqkm
FROM forestation
WHERE year = 2016 AND
    (total_area_sq_mi * 2.59) <=
      /* subquery to extract difference in area */
      (
  WITH yr1990 AS
  (
      SELECT forest_area_sqkm AS area1990, year, country_code
      FROM forestation
      WHERE year = '1990' AND country_name = 'World'
  ),
  yr2016 AS
  (
      SELECT forest_area_sqkm AS area2016, year, country_code
      FROM forestation
      WHERE year = '2016' AND country_name = 'World'
  )
  SELECT (area1990 - area2016)
  FROM yr1990
  JOIN yr2016
  ON yr1990.country_code = yr2016.country_code)
ORDER BY sqkm DESC
LIMIT 1;
```

2. REGIONAL OUTLOOK

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

```
CREATE TABLE percent_forest_per_region AS
SELECT region,
ROUND(CAST((100*SUM(forest_area_sqkm)/(SUM(total_area_sq_mi)*2.59)) AS
numeric), 2) as percent_land, year
FROM forestation
WHERE year = '2016' OR year = '1990'
GROUP BY 1, 3
ORDER BY percent_land 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 percent_land
FROM percent_forest_per_region
WHERE year = '2016' AND region = 'World';
```

```
SELECT region, percent_land
FROM percent_forest_per_region
WHERE year = '2016'
ORDER BY percent_land DESC
LIMIT 1;
```

```
SELECT region, percent_land
FROM percent_forest_per_region
WHERE year = '2016'
ORDER BY percent_land
LIMIT 1;
```

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 percent_land
FROM percent_forest_per_region
WHERE year = '1990' AND region = 'World';
```

```
SELECT region, percent_land
FROM percent_forest_per_region
```

```
WHERE year = '1990'
ORDER BY percent_land DESC
LIMIT 1;
```

```
SELECT region, percent_land
FROM percent_forest_per_region
WHERE year = '1990'
ORDER BY percent_land
LIMIT 1;
```

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

```
WITH yr1990 AS  (
      SELECT region, percent_land AS land1990
      FROM percent_forest_per_region
      WHERE year = '1990'
),
yr2016 AS
(
      SELECT region, percent_land AS land2016
      FROM percent_forest_per_region
      WHERE year = '2016'
)
SELECT *
FROM yr1990
JOIN yr2016
ON yr1990.region = yr2016.region
WHERE (land1990 - land2016) > 0;
```

3. COUNTRY-LEVEL DETAIL

Instructions:

- Answering these questions will help you add information into the template.
- Use these questions as guides to write SQL queries.
- Use the output from the query to answer these questions.

```
/* area increased: absolute forest area change by country, top 5 */
WITH yr1990 AS  (
      SELECT country_name, region, forest_area_sqkm AS forest1990
      FROM forestation
      WHERE year = '1990' AND country_name != 'World'
),
```

```
yr2016 AS
(
      SELECT country_name, region, forest_area_sqkm AS forest2016
      FROM forestation
      WHERE year = '2016' AND country_name != 'World'
)
SELECT *,
      ROUND(CAST((forest2016 - forest1990) AS numeric), 2) AS
absolute_forest_area_change
FROM yr1990
JOIN yr2016
ON yr1990.country_name = yr2016.country_name
WHERE forest1990 IS NOT NULL AND forest2016 IS NOT NULL
ORDER BY absolute_forest_area_change DESC
LIMIT 5;
```

```
/* area increased: percent forest area change by country, top 1 */
WITH yr1990 AS  (
      SELECT country_name, region, forest_area_sqkm AS forest1990
      FROM forestation
      WHERE year = '1990' AND country_name != 'World'
),
yr2016 AS
(
      SELECT country_name, region, forest_area_sqkm AS forest2016
      FROM forestation
      WHERE year = '2016' AND country_name != 'World'
)
SELECT *,
      ROUND(CAST(((forest2016 - forest1990)/forest1990) * 100 AS numeric),
2) AS percent_forest_area_change
FROM yr1990
JOIN yr2016
ON yr1990.country_name = yr2016.country_name
WHERE forest1990 IS NOT NULL AND forest2016 IS NOT NULL
ORDER BY percent_forest_area_change DESC
LIMIT 1;
```

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?

```
/* absolute forest area change by country, top 5 */
```

```
WITH yr1990 AS  (
      SELECT country_name, region, forest_area_sqkm AS forest1990
      FROM forestation
      WHERE year = '1990' AND country_name != 'World'
),
yr2016 AS
(
      SELECT country_name, region, forest_area_sqkm AS forest2016
      FROM forestation
      WHERE year = '2016' AND country_name != 'World'
)
SELECT *,
    ROUND(CAST((forest1990 - forest2016) AS numeric), 2) AS
absolute_forest_area_change
FROM yr1990
JOIN yr2016
ON yr1990.country_name = yr2016.country_name
WHERE forest1990 > forest2016
ORDER BY absolute_forest_area_change 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?

```
/* area decrease: percent forest area change by country, top 5 */
WITH yr1990 AS  (
      SELECT country_name, region, forest_area_sqkm AS forest1990
      FROM forestation
      WHERE year = '1990' AND country_name != 'World'
),
yr2016 AS
(
      SELECT country_name, region, forest_area_sqkm AS forest2016
      FROM forestation
      WHERE year = '2016' AND country_name != 'World'
)
SELECT *,
      ROUND(CAST(((forest2016 - forest1990)/forest1990) * 100 AS numeric),
2) AS percent_forest_area_change
FROM yr1990
JOIN yr2016
```

```
ON yr1990.country_name = yr2016.country_name
WHERE forest1990 IS NOT NULL AND forest2016 IS NOT NULL
ORDER BY percent_forest_area_change
LIMIT 5;
```

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

```
/* count # of countries in each quartile in 2016 */
SELECT DISTINCT(quartile) AS quartile, COUNT(country_name) OVER (PARTITION
BY quartile) AS number_of_countries
FROM (
  SELECT country_name,
      CASE
      WHEN percent_designated_as_forest <= 25 THEN '1st'
      WHEN percent_designated_as_forest > 25 AND
percent_designated_as_forest <= 50 THEN '2nd'
      WHEN percent_designated_as_forest > 50 AND
  percent_designated_as_forest <= 75 THEN '3rd'
      ELSE '4th'
      END AS quartile
  FROM forestation
  WHERE year = '2016' AND country_name != 'World' AND
percent_designated_as_forest IS NOT NULL) AS subquery;
```

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

```
/* count countries with percent forestation > 75 in 2016 */
SELECT COUNT(*) AS num_of_countries, country_name, region,
ROUND(CAST(percent_designated_as_forest AS numeric), 2)
FROM forestation
WHERE year = '2016' AND country_name != 'World' AND
percent_designated_as_forest IS NOT NULL AND
percent_designated_as_forest > 75.
GROUP BY 2, 3, 4
ORDER BY 4 DESC;
```

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

```
/* count countries with higher forestation % than US in 2016 */
SELECT COUNT(*) AS num_of_countries, country_name,
percent_designated_as_forest
```

```
FROM forestation
WHERE year = '2016' AND country_name != 'World' AND
percent_designated_as_forest IS NOT NULL AND
percent_designated_as_forest > (
  SELECT percent_designated_as_forest
  FROM forestation
  WHERE year = '2016' AND country_name = 'United States')
GROUP BY 2, 3
ORDER BY 3;
```

94 countries had a percent forestation higher than the United States in 2016.