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 82016472.036028 sq km in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 79825433.9505107 sq km, a loss of 2191038.08551738 sq km, or 2.67%.

The forest area lost over this time period is slightly more than the entire land area of Saudi Arabia listed for the year 2016 (which is 2149690 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 51.02%, 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	51.0299798667514%	46.1620721996047%
Europe & Central Asia	37.2839398564019%	38.0414216032517%
North America	35.6511790009015%	36.0393609681438%

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 527229.062. 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, much lower than the figure for China.

India and Russian Federation 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. China increased in forest area by 33.55% 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
Indonesia	East Asia & Pacific	282193.9844
Myanmar	East Asia & Pacific	107234.0039
Nigeria	Sub-Saharan Africa	106506.00098
Tanzania	Sub-Saharan Africa	102320

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
4 (100 - 75)	9
3 (75 - 50)	38
2 (50 - 25)	73
1 (25 - 0)	98

The largest number of countries in 2016 were found in the 1 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	98.2576939676578%
Micronesia, Fed. Sts.	East Asia & Pacific	91.8572390715248%
Gabon	Sub-Saharan Africa	90.0376418700565%
Seychelles	Sub-Saharan Africa	88.4111367385789%
Palau	East Asia & Pacific	87.6068085491204%
American Samoa	East Asia & Pacific	87.5000875000875%
Guyana	Latin America & Caribbean	83.9014489110682%
Lao PDR	East Asia & Pacific	82.1082317640861%
Solomon Islands	East Asia & Pacific	77.8635177945066%

5. RECOMMENDATIONS

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

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

After analyzing the World Bank data, I concluded that the total forest area of the world has decreased from 82016472.036028 sq km 79825433.9505107 sq km just within 26 years. The regions that significantly impacted the worlds record of forest area were Latin America & Caribbean (dropped from 51.03% to 46.16%) specifically Brazil and Sub-Saharan Africa (30.67% to 28.79%) and countries such as Togo, Nigeria, and Uganda. The rest of the regions had an increase in forest area over that period of time. However, there is one country that did very well in increasing the forest area, China. From 1990 to 2016 it had increased its forest area by 1.34%. I also learned that only 9 countries in 2016 had forest area above 75% to the total area of a country. Most of these countries are in Latin America & Caribbean, Sub-Saharan Africa, and East Asia & Pacific.

The region that is most affected is Sub-Saharan Africa hence I looked into each country. Most of the countries in Sub-Saharan Africa are low income and the Top 3 countries that had the largest absolute difference between 1990 and 2016 in forest area are Nigeria, Tanzania and Zimbabwe. Three countries combined make about 292970 km sq loss in forest area. However, there were 8 countries in the Sub-Saharan Africa region which gained some forest area by 2016. The sum in absolute difference is 24951 km sq. This amount of forest area gain is very insignificant compared to what was lost in the Top 3 countries. Hence the Sub-Saharan Africa region is one of the top that made it to the top of the list of decrease in forest area.

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

I think we should focus on countries such as Nigeria. It is not only one of the countries in the Top 5 list of decrease in percent (61.80%) but also absolute difference in forest area (106506.00098). We should also look into how China was able to bring their forest area by 1.34% within 26 years. I would also focus on small countries like Togo where the total area is just 21000 km sq out of which the forest area has come down from 6850 to 1682 km sq, because it is also a low-income area I would consider understanding why the area of forest is rapidly decreasing.

6. Appendix: SQL queries used

```
-- Steps to Complete
-- Create a View called “forestation” by joining all three tables - forest_area, land_area and regions in the workspace.
-- The forest_area and land_area tables join on both country_code AND year.
-- The regions table joins these based on only country_code.
-- 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.
-- 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).

```
DROP VIEW forestation;
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,  100*(f.forest_area_sqkm/(l.total_area_sq_mi*2.59))
forest_percent
  FROM forest_area f
  JOIN land_area l
  ON f.year = l.year AND f.country_code = l.country_code
  JOIN regions r
  ON r.country_code = l.country_code);
```

-- 1. GLOBAL SITUATION
-- 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.

-- 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.
-- answer: 82016472.036028

```
SELECT SUM(forest_area_sqkm) sum_area_1990
FROM forestation
WHERE year = 1990
GROUP BY year
```

-- 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.”
-- answer: 79825433.9505107

```
SELECT SUM(forest_area_sqkm) sum_area_1990
FROM forestation
```

```
WHERE year = 2016
GROUP BY year

-- c. What was the change (in sq km) in the forest area of the world from 1990 to 2016?
-- answer: 2191038.08551738

SELECT t1.sum_area, t1.year, LAG(t1.sum_area) OVER ( ORDER BY t1.year) lag,
(LAG(t1.sum_area) OVER ( ORDER BY t1.year) - t1.sum_area) diff_area,

FROM (SELECT SUM(forest_area_sqkm) sum_area, year
FROM forestation
WHERE year = 1990 OR year = 2016
GROUP BY 2
ORDER BY year) t1

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

SELECT t1.sum_area, t1.year, LAG(t1.sum_area) OVER ( ORDER BY t1.year) lag,
(LAG(t1.sum_area) OVER ( ORDER BY t1.year) - t1.sum_area) diff_area,
100*( (LAG(t1.sum_area) OVER ( ORDER BY t1.year) - t1.sum_area)/LAG(t1.sum_area) OVER
( ORDER BY t1.year)) per_change

FROM (SELECT SUM(forest_area_sqkm) sum_area, year
FROM forestation
WHERE year = 1990 OR year = 2016
GROUP BY 2
ORDER BY year) t1

-- 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, year, forest_area_sqkm
FROM forest_area
WHERE year = 2016 AND (forest_area_sqkm >= 1999999 AND forest_area_sqkm <=
(SELECT (LAG(t1.sum_area) OVER ( ORDER BY t1.year) - t1.sum_area) diff_area
FROM (SELECT SUM(forest_area_sqkm) sum_area, year
FROM forest_area
WHERE year = 1990 OR year = 2016
GROUP BY 2
ORDER BY year) t1
ORDER BY diff_area
LIMIT 1))
```

-- 2. REGIONAL OUTLOOK
-- 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.

-- 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).
-- Based on the table you created, ...

-- a. What was the percent forest of the entire world in 2016?

-- answer: 31.38

```
SELECT 100*(SUM(forest_area_sqkm)/(SUM(total_area_sq_mi*2.59))) sum_percent, region, year
FROM forestation
WHERE year = 2016 AND region = 'World'
GROUP BY 2 , 3
ORDER BY 1
```

-- Which region had the HIGHEST percent forest in 2016
-- answer: 46.1620721996047 Latin America & Caribbean

```
SELECT 100*(SUM(forest_area_sqkm)/(SUM(total_area_sq_mi*2.59))) sum_percent, region, year
FROM forestation
WHERE year = 2016
GROUP BY 2 , 3
ORDER BY 1 DESC
LIMIT 1
```

-- and which had the LOWEST, to 2 decimal places?
-- answer:2.06826486871501Middle East & North Africa

```
SELECT CAST(100*(SUM(forest_area_sqkm)/(SUM(total_area_sq_mi*2.59))) as DECIMAL(10,2)) sum_percent, region, year
FROM forestation
WHERE year = 2016
GROUP BY 2 , 3
```

```
ORDER BY 1
LIMIT 1


-- b. What was the percent forest of the entire world in 1990?
-- answer: 32.4222035575689

SELECT  100*(SUM(forest_area_sqkm)/(SUM(total_area_sq_mi*2.59))) sum_percent, region,
year
FROM forestation
WHERE year = 1990 AND region = 'World'
GROUP BY 2 , 3
ORDER BY 1


-- Which region had the HIGHEST percent forest in 1990,
-- answer: 51.0299798667514          Latin America & Caribbean

SELECT          CAST(100*(SUM(forest_area_sqkm)/(SUM(total_area_sq_mi*2.59)))          as
DECIMAL(10,2)) sum_percent, region, year
FROM forestation
WHERE year = 1990
GROUP BY 2 , 3
ORDER BY 1 DESC
LIMIT 1

-- and which had the LOWEST, to 2 decimal places?
-- answer: 1.77524062469353          Middle East & North Africa

SELECT          CAST(100*(SUM(forest_area_sqkm)/(SUM(total_area_sq_mi*2.59)))          as
DECIMAL(10,2)) sum_percent, region, year
FROM forestation
WHERE year = 1990
GROUP BY 2 , 3
ORDER BY 1
LIMIT 1


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

```
WITH t90 AS (SELECT SUM(forest_area_sqkm) fa_sum, CAST(100*(SUM(forest_area_sqkm)/(SUM(total_area_sq_mi*2.59))) as DECIMAL(10,2)) sum_percent, region, year
FROM forestation
WHERE year = 1990
GROUP BY 3 , 4
ORDER BY 1 ),

t16 AS (SELECT SUM(forest_area_sqkm) fa_sum, CAST(100*(SUM(forest_area_sqkm)/(SUM(total_area_sq_mi*2.59))) as DECIMAL(10,2)) sum_percent, region, year
FROM forestation
WHERE year = 2016
GROUP BY 3 , 4
ORDER BY 1 )

SELECT t16.region region, t16.fa_sum fa_sum_16, t16.sum_percent fa_percent_16, t90.fa_sum fa_sum_1990, t90.sum_percent fa_percent_90, (t90.fa_sum - t16.fa_sum) diff_FA
FROM t90
JOIN t16
ON t16.region = t90.region
WHERE (t90.fa_sum - t16.fa_sum) > 0
ORDER BY diff_FA

-- 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.

-- a. Which 5 countries saw the largest amount increase in forest area from 1990 to 2016?
-- What was the difference in forest area for each?

WITH t90 AS (SELECT SUM(forest_area_sqkm) fa_sum, CAST(100*(SUM(forest_area_sqkm)/(SUM(total_area_sq_mi*2.59))) as DECIMAL(10,2)) sum_percent, country_name, year
FROM forestation
WHERE year = 1990
GROUP BY 3 , 4
ORDER BY 1 ),
```

```
t16 AS (SELECT SUM(forest_area_sqkm) fa_sum,
CAST(100*(SUM(forest_area_sqkm)/(SUM(total_area_sq_mi*2.59))) as DECIMAL(10,2))
sum_percent, country_name, year
FROM forestation
WHERE year = 2016
GROUP BY 3 , 4
ORDER BY 1 )

SELECT t16.country_name country_name, t16.fa_sum fa_sum_16, t16.sum_percent
fa_percent_16, t90.fa_sum fa_sum_1990, t90.sum_percent fa_percent_90, -(t90.fa_sum -
t16.fa_sum) diff_FA, t16.fa_sum/t90.fa_sum ratio_inc
FROM t90
JOIN t16
ON t16.country_name = t90.country_name
WHERE (t90.fa_sum - t16.fa_sum) < 0
ORDER BY diff_FA 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 t90 AS (SELECT SUM(forest_area_sqkm) fa_sum,
CAST(100*(SUM(forest_area_sqkm)/(SUM(total_area_sq_mi*2.59))) as DECIMAL(10,2))
sum_percent, country_name, region, year
FROM forestation
WHERE year = 1990
GROUP BY 3 , 4, 5
ORDER BY 1 ),

t16 AS (SELECT SUM(forest_area_sqkm) fa_sum,
CAST(100*(SUM(forest_area_sqkm)/(SUM(total_area_sq_mi*2.59))) as DECIMAL(10,2))
sum_percent, country_name, region, year
FROM forestation
WHERE year = 2016
GROUP BY 3 , 4 ,5
ORDER BY 1 )

SELECT t16.region region, t16.country_name country_name, t16.fa_sum fa_sum_16,
t16.sum_percent fa_percent_16, t90.fa_sum fa_sum_1990, t90.sum_percent
fa_percent_90, (t90.fa_sum - t16.fa_sum) diff_FA, CAST(100*(1 - t16.fa_sum/t90.fa_sum) AS
DECIMAL(10,2)) ratio_inc
FROM t90
JOIN t16
```

```
ON t16.country_name = t90.country_name
WHERE (t90.fa_sum - t16.fa_sum) > 0
ORDER BY ratio_inc DESC
LIMIT 5

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

WITH t1 AS (SELECT
CASE
WHEN forest_percent > 75 THEN 4
WHEN forest_percent <= 75 AND forest_percent > 50 THEN 3
WHEN forest_percent <= 50 AND forest_percent > 25 THEN 2
WHEN forest_percent <= 25 THEN 1 END percent_group, year, forest_percent, country_name
FROM forestation)

SELECT t1.percent_group, COUNT(t1.percent_group) num_group
FROM t1
WHERE t1.year = 2016
GROUP BY 1

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

WITH t1 AS (SELECT
CASE
WHEN forest_percent > 75 THEN 4
WHEN forest_percent <= 75 AND forest_percent > 50 THEN 3
WHEN forest_percent <= 50 AND forest_percent > 25 THEN 2
WHEN forest_percent <= 25 THEN 1 END percent_group, year, forest_percent, country_name,
region
FROM forestation)

SELECT t1.country_name, t1.region,  t1.forest_percent
FROM t1
WHERE t1.year = 2016 AND t1.percent_group = 4
ORDER BY 3 DESC


-- e. How many countries had a percent forestation higher than the United States in 2016?
-- answer: 94
SELECT COUNT(*)
FROM forestation
WHERE year = 2016 AND forest_percent >
```

```
(SELECT forest_percent
FROM forestation
WHERE country_name = 'United States' AND year = 2016)

-- Some additional SQL Queries



WITH t16 AS(SELECT SUM(forest_area_sqkm) fa_sum, country_name, year, income_group
FROM forestation
WHERE (year = 2016) AND region = 'Sub-Saharan Africa'
GROUP BY 2 , 3 ,4
ORDER BY year DESC),

t90 AS(SELECT SUM(forest_area_sqkm) fa_sum, country_name, year, income_group
FROM forestation
WHERE (year = 1990) AND region = 'Sub-Saharan Africa'
GROUP BY 2 , 3 ,4
ORDER BY year DESC)

SELECT (t90.fa_sum - t16.fa_sum) abs_difference, t16.country_name, t16.income_group
income_in_2016, t90.income_group income_in_1990
FROM t16
JOIN t90
ON t16.country_name = t90.country_name
ORDER BY 1 DESC
```