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 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 1,279,999.9891 sqkm).

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%
Sub-Saharan Africa	32.19%	27.56%
South Asia	16.51%	17.50%

North America	35.65%	36.04%
Middle East & North Africa	1.78%	2.07%
Latin America & Caribbean	51.03%	46.16%
Europe & Central Asia	37.27%	38.06%
East Asia & Pacific	25.77%	26.36%

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 (32.19% to 27.56%). 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 527,229.06 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 of America, but it only saw an increase of 79,200 sqkm, much lower than the figure for China.

China and USA 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 213.67% 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	-541,510.00 sqkm

Indonesia	East Asia & Pacific	-282,193.98 sqkm
Myanmar	East Asia & Pacific	-107,234.00 sqkm
Nigeria	Sub-Saharan Africa	-106,506.00 sqkm
Tanzania	Sub-Saharan Africa	-102,320.00 sqkm

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	85

2	72
3	38
4	9

The largest number of countries in 2016 were found in the First (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	% 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

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's total forest cover has been perpetually decreasing over the past 26 years, since the data collection started in 1990.
- The Region which had the greatest forest cover w.r.t to their land area in 1990 (Latin America & Caribbean) was also the one which lost the most forest cover w.r.t the land area by 2016 (4.87%).
- The income levels of the countries seem to be playing a role towards deforestation around the world.
    - The low-income and lower-middle income countries not only have the lowest relative forest cover w.r.t their land area, but also have lost forest areas to deforestation since 1990. The low-income countries have had a percentage change of -15.21% in their forest area while the lower-middle income countries have had a percentage change of -1.30% in their forest area.

- The High-Income countries on the other hand are performing really well in terms of forest cover. The upper-middle income and high-income countries not only have greater relative forest cover w.r.t their land area but also have increased their forest cover since 1990.
- It may be concluded that the better the income levels of a country, the less prone to deforestation that country may be.

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

The countries in Sub-Saharan Africa should be of special focus.
Most countries in this region fall in either the low income or lower-middle income category (38 out of 45 countries). Counties in this region, therefore, are more prone to deforestation.

Togo, Nigeria and Uganda are in need of urgent attention as they have lost over half of their forest area since 1990.
Nigeria in particular requires attention as it also ranks among the countries losing the largest forested area since 1990.

APPENDIX (SQL Code)

Creating the VIEW "forestation"

```
CREATE VIEW forestation AS
SELECT f.country_code c_code, f.country_name country, r.region, f.year yr, f.forest_area_sqkm
forest_area, l.total_area_sq_mi*2.59 total_area,
r.income_group, f.forest_area_sqkm/(total_area_sq_mi*2.59)*100 percent_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 l.country_code = r.country_code
ORDER BY 1,4

CREATE VIEW forestation_2 AS
SELECT *
FROM forestation
WHERE percent_forest IS NOT NULL
```

//Global Situation//

(a) What was the total forest area (in sq km) of the world in 1990? and
(b) What was the total forest area (in sq km) of the world in 2016?

```
SELECT region, yr, forest_area, total_area, percent_forest
FROM forestation_2
WHERE region LIKE 'Worl%'
AND (yr = '1990' OR yr = '2016')
```

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

```
WITH t1 AS
  (SELECT region, yr, forest_area, total_area, percent_forest
FROM forestation_2
WHERE region LIKE 'Worl%'
AND (yr = '1990' OR yr = '2016'))

SELECT MAX(forest_area) - MIN(forest_area) change_forest_area
FROM t1
```

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

```
WITH t1 AS (SELECT *, LAG(forest_area) OVER(PARTITION BY country ORDER BY yr)
AS forest_area_1990
FROM forestation_2
WHERE yr = '1990' OR yr = '2016'
```

```
ORDER BY 1,4),

t2 AS (SELECT *, forest_area - forest_area_1990 AS area_change,
   (forest_area - forest_area_1990)/forest_area_1990*100 AS percent_change
FROM t1
WHERE forest_area_1990 IS NOT NULL)

SELECT country, percent_change
FROM t2
WHERE country LIKE 'Worl%'
```

(e)If you compare the amount of forest area lost between 1990 and 2016, to which countrys total area in 2016 is it closest to

```
SELECT country, total_area
FROM forestation_2
WHERE total_area < 1324449 AND yr = '2016'
ORDER BY 2 DESC
LIMIT 1
```

//Regional Situation//

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.

```
SELECT region, yr, SUM(forest_area)/SUM(total_area)*100 regional_forest_cover
FROM forestation_2
WHERE (yr ='1990' OR yr = '2016')
GROUP BY 1,2
ORDER BY 1,2
```

(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?

```
WITH t1 AS (
SELECT region, yr, SUM(forest_area)/SUM(total_area)*100 regional_forest_cover
FROM forestation_2
GROUP BY 1,2
ORDER BY 1,2)

SELECT region, yr, regional_forest_cover
FROM t1
WHERE yr = '2016'
ORDER BY 3 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 t1 AS (
SELECT region, yr, SUM(forest_area)/SUM(total_area)*100 regional_forest_cover
FROM forestation_2
GROUP BY 1,2
ORDER BY 1,2)

SELECT region, yr, regional_forest_cover
FROM t1
WHERE yr = '1990'
ORDER BY 3 DESC
```

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

```
WITH t1 AS (
SELECT region, yr, SUM(forest_area)/SUM(total_area)*100 regional_forest_cover
FROM forestation_2
WHERE yr = '1990' OR yr = '2016'
GROUP BY 1,2
ORDER BY 1,2),

t2 AS (SELECT region, yr, regional_forest_cover,
LAG(regional_forest_cover) OVER(PARTITION BY region ORDER BY yr)
FROM t1),

t3 AS (SELECT *,
CASE WHEN (regional_forest_cover-lag) < 0 THEN 'DECREASE' END AS forest_cover
FROM t2)

SELECT *
FROM t3
WHERE forest_cover IS NOT NULL
```

##OR##

```
WITH t1 AS (
SELECT region, yr, SUM(forest_area)/SUM(total_area)*100 regional_forest_cover
FROM forestation_2
GROUP BY 1,2
ORDER BY 1,2)

SELECT *
FROM t1 a
JOIN t1 b
ON a.region = b.region
WHERE (a.yr = '1990' AND b.yr = '2016')
AND (a.regional_forest_cover > b.regional_forest_cover)
```

//Country-Wise Situation

(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 t1 AS
(SELECT country, region, yr, forest_area forest_cover_area
FROM forestation_2
ORDER BY 1,3),

t2 AS
(SELECT *, LAG(forest_cover_area) OVER(PARTITION BY country ORDER BY yr),
forest_cover_area - LAG(forest_cover_area) OVER(PARTITION BY country ORDER BY yr)
AS diff_area
FROM t1)

SELECT country, region, SUM(diff_area)
FROM t2
WHERE country NOT LIKE 'World'
GROUP BY 1,2
ORDER BY 3
LIMIT 5
```

(b)Which 5 countries saw the largest percent decrease in

```
WITH t1 AS (SELECT *, LAG(forest_area) OVER(PARTITION BY country ORDER BY yr)
AS forest_area_1990
FROM forestation_2
WHERE yr = '1990' OR yr = '2016'
ORDER BY 1,4),

t2 AS (SELECT *, forest_area - forest_area_1990 AS area_change,
   (forest_area - forest_area_1990)/forest_area_1990*100 AS percent_change
FROM t1
WHERE forest_area_1990 IS NOT NULL)

SELECT country, region, percent_change
FROM t2
WHERE country NOT LIKE 'World'
ORDER BY 3
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 (percent_forest <= 25) THEN 1
   WHEN (percent_forest > 25 AND percent_forest <= 50) THEN 2
```

```
    WHEN (percent_forest > 50 AND percent_forest <= 75) THEN 3
    WHEN (percent_forest > 75 AND percent_forest <= 100) THEN 4 END AS quartile
FROM forestation_2)

SELECT quartile, COUNT(*)
FROM t1
WHERE yr = '2016' AND country NOT LIKE 'World'
GROUP BY 1
ORDER 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 (percent_forest <= 25) THEN 1
    WHEN (percent_forest > 25 AND percent_forest <= 50) THEN 2
    WHEN (percent_forest > 50 AND percent_forest <= 75) THEN 3
    WHEN (percent_forest > 75 AND percent_forest <= 100) THEN 4 END AS quartile
FROM forestation_2),

t2 AS (SELECT quartile, COUNT(*)
FROM t1
WHERE yr = '2016' AND country NOT LIKE 'World'
GROUP BY 1
ORDER BY 1)

SELECT country, region, percent_forest
FROM t1
WHERE quartile = 4 AND yr = '2016'
ORDER BY 3 DESC
```

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

```
SELECT COUNT(*)
FROM forestation_2
WHERE percent_forest > (SELECT percent_forest
        FROM forestation_2
        WHERE country LIKE 'United States%' AND yr = '2016')
AND yr = '2016'
```

// Recommendations

```
WITH t1 AS (SELECT *, LAG(forest_area) OVER(PARTITION BY country ORDER BY yr)
AS forest_area_1990
FROM forestation_2
WHERE yr = '1990' OR yr = '2016'
ORDER BY 1,4),

t2 AS (SELECT *, forest_area - forest_area_1990 AS area_change,
```

```
   (forest_area - forest_area_1990)/forest_area_1990*100 AS percent_change
FROM t1
WHERE forest_area_1990 IS NOT NULL)

SELECT income_group, AVG(percent_change) AVG_CHANGE, AVG(percent_forest)
AVG_FOREST
FROM t2
WHERE income_group NOT LIKE NULL
GROUP BY 1
ORDER BY 2,3

##
##

WITH t1 AS (SELECT *, LAG(forest_area) OVER(PARTITION BY country ORDER BY yr)
AS forest_area_1990
FROM forestation_2
WHERE yr = '1990' OR yr = '2016'
ORDER BY 1,4),

t2 AS (SELECT *, forest_area - forest_area_1990 AS area_change,
   (forest_area - forest_area_1990)/forest_area_1990*100 AS percent_change
FROM t1
WHERE forest_area_1990 IS NOT NULL)

SELECT region, income_group, COUNT(*)
FROM t2
WHERE income_group NOT LIKE 'NULL' AND region LIKE 'Sub%'
GROUP BY 1, 2
ORDER BY 3 DESC
```