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 41282694.9 in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 39958245.9, a loss of 1324449, 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).

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
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**. 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**, much lower than the figure for **Iceland**.

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:

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:

Quadfdrtile	Number of Countries
0-25%	85
25-50%	72
50-75%	38
75-100%	9

The largest number of countries in 2016 were found in the **0-25%** 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
American Samoa	Sub-Saharan Africa	98.26
Micronesia, Fed. Sts	East Asia & Pacific	91.86
Gabon	Sub-Saharan Africa	90.04
Guyana	Latin America & Caribbean	83.90
Lao PDR	East Asia & Pacific	82.11
Palau	East Asia & Pacific	87.61

Solomon Islands	East Asia & Pacific	77.86
Suriname	Latin America & Caribbean	98.26
Seychelles	Sub-Saharan Africa	88.41

4. 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?*

Deforestation has become increasingly serious from 1990 to 2016. We should primarily focus on countries in the Sub-Saharan Africa Region.

5. APPENDIX: SQL queries used

```sql
-- Create a View called "forestation" by joining all three tables - forest_area,
land_area and regions in the workspace.
DROP VIEW IF EXISTS forestation;

CREATE VIEW forestation AS
SELECT f.country_code country_code,
       f.country_name country_name,
       f.year "year",
       f.forest_area_sqkm forest_area_sqkm,
       l.total_area_sq_mi total_area_sq_mi,
       l.total_area_sq_mi * 2.59 total_area_sqkm,
       r.region region,
       r.income_group income_group,

       (f.forest_area_sqkm/(l.total_area_sq_mi*2.59))*100 pct_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;

SELECT * FROM forestation;

-------------------------- 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.

DROP VIEW IF EXISTS global_situation;
```

```sql
CREATE VIEW global_situation AS
WITH t_2016 AS (
    SELECT year, region, forest_area_sqkm
    FROM forestation
    WHERE region = 'World' AND year = 2016),
t_1990 AS (
    SELECT year, region, forest_area_sqkm
    FROM forestation
    WHERE region = 'World' AND year = 1990)
SELECT t_1990.region,
       t_1990.forest_area_sqkm forest_area_sqkm_1990,
       t_2016.forest_area_sqkm forest_area_sqkm_2016
FROM t_1990, t_2016;


------------------------------------------------------------------
-- 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 year, region, forest_area_sqkm
FROM forestation
WHERE region = 'World' AND year = 1990;

-- year     region forest_area_sqkm
-- 1990    World  41282694.9


------------------------------------------------------------------
-- 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 year, region, forest_area_sqkm
FROM forestation
WHERE region = 'World' AND year = 2016;

-- year     region forest_area_sqkm
-- 2016    World  39958245.9


------------------------------------------------------------------
-- c. What was the change (in sq km) in the forest area of the world from 1990 to
2016?
SELECT forest_area_sqkm_2016 - forest_area_sqkm_1990 forest_area_change_sqkm
FROM global_situation;

-- forest_area_change_sqkm
-- -1324449


------------------------------------------------------------------
-- d. What was the percent change in forest area of the world between 1990 and 2016?
SELECT (forest_area_sqkm_2016 - forest_area_sqkm_1990)/forest_area_sqkm_1990 *100
forest_area_pct_change_sqkm
FROM global_situation;

-- forest_area_pct_change_sqkm
-- -3.20824258980244


------------------------------------------------------------------
-- 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 f.country_name,
       f.total_area_sqkm,
       ABS(f.total_area_sqkm - ( g.forest_area_sqkm_1990 - g.forest_area_sqkm_2016))
diff_area
FROM forestation f, global_situation g
ORDER BY diff_area ASC
```

```sql
LIMIT 1;

-- country_name    total_area_sqkm    diff_area
-- Peru     1279999.9891    44449.0109000001


-------------------------- 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, ....

DROP VIEW IF EXISTS regional_outlook;

CREATE VIEW regional_outlook AS
WITH t_2016 AS (
    SELECT  region, ROUND(CAST(SUM(forest_area_sqkm)/SUM(total_area_sqkm)*100 AS
NUMERIC), 2) pct_forest_area_2016
    FROM forestation
    WHERE year = 2016
    GROUP BY region),
t_1990 AS (
    SELECT  region, ROUND(CAST(SUM(forest_area_sqkm)/SUM(total_area_sqkm)*100 AS
NUMERIC), 2) pct_forest_area_1990
    FROM forestation
    WHERE year = 1990
    GROUP BY region)
SELECT t_1990.region, t_1990.pct_forest_area_1990, t_2016.pct_forest_area_2016
FROM t_1990
JOIN t_2016
ON t_1990.region = t_2016.region;

SELECT * FROM regional_outlook;

-- region  pct_forest_area_1990   pct_forest_area_2016
-- Latin America & Caribbean   51.03  46.16
-- Sub-Saharan Africa  30.67  28.79
-- Europe & Central Asia   37.28  38.04
-- East Asia & Pacific 25.78  26.36
-- South Asia  16.51  17.51
-- Middle East & North Africa  1.78   2.07
-- World   32.42  31.38
-- North America   35.65  36.04


-----------------------------------------------------------------
-- 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 region, pct_forest_area_2016
FROM regional_outlook
WHERE region = 'World';

-- region  pct_forest_area_2016
-- World   31.38

SELECT region, pct_forest_area_2016
FROM regional_outlook
```

```sql
WHERE region != 'World'
ORDER BY pct_forest_area_2016 DESC
LIMIT 1;

-- region    pct_forest_area_2016
-- Latin America & Caribbean    46.16

SELECT region, pct_forest_area_2016
FROM regional_outlook
WHERE region != 'World'
ORDER BY pct_forest_area_2016 ASC
LIMIT 1;

-- region    pct_forest_area_2016
-- Middle East & North Africa  2.07


-------------------------------------------------------------------
-- 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 region, pct_forest_area_1990
FROM regional_outlook
WHERE region = 'World'

-- region    pct_forest_area_1990
-- World    32.42

SELECT region, pct_forest_area_1990
FROM regional_outlook
WHERE region != 'World'
ORDER BY pct_forest_area_1990 DESC
LIMIT 1;

-- region    pct_forest_area_1990
-- Latin America & Caribbean    51.03

SELECT region, pct_forest_area_1990
FROM regional_outlook
WHERE region != 'World'
ORDER BY pct_forest_area_1990 ASC
LIMIT 1;

-- region    pct_forest_area_1990
-- Middle East & North Africa  1.78


-------------------------------------------------------------------
-- c. Based on the table you created, which regions of the world DECREASED in forest
area from 1990 to 2016?
SELECT region, pct_forest_area_1990, pct_forest_area_2016
FROM regional_outlook;

-- region  pct_forest_area_1990   pct_forest_area_2016
-- Latin America & Caribbean    51.03  46.16
-- Sub-Saharan Africa 30.67  28.79
-- Europe & Central Asia   37.28  38.04
-- East Asia & Pacific 25.78  26.36
-- South Asia  16.51  17.51
-- Middle East & North Africa  1.78   2.07
-- World   32.42  31.38
-- North America   35.65  36.04
```

```sql
SELECT region, pct_forest_area_1990, pct_forest_area_2016
FROM regional_outlook
WHERE pct_forest_area_1990 > pct_forest_area_2016;

-- region  pct_forest_area_1990   pct_forest_area_2016
-- Latin America & Caribbean   51.03  46.16
-- Sub-Saharan Africa  30.67  28.79
-- World   32.42  31.38


------------------------ 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.

DROP VIEW IF EXISTS country_detail;

CREATE VIEW country_detail AS
WITH t_2016 AS (
    SELECT country_name, forest_area_sqkm
    FROM forestation
    WHERE year = 2016 AND forest_area_sqkm IS NOT NULL),
t_1990 AS (
    SELECT country_name, region, forest_area_sqkm
    FROM forestation
    WHERE year = 1990 AND forest_area_sqkm IS NOT NULL)
SELECT t_1990.country_name,
       t_1990.region,
       t_1990.forest_area_sqkm forest_area_sqkm_1990,
       t_2016.forest_area_sqkm forest_area_sqkm_2016
FROM t_1990
JOIN t_2016
ON t_1990.country_name = t_2016.country_name;

SELECT * FROM country_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?

--- SUCCESS STORIES
SELECT country_name, forest_area_sqkm_2016 - forest_area_sqkm_1990 diff_forest_area
FROM country_detail
WHERE country_name != 'World'
ORDER BY diff_forest_area DESC
LIMIT 5;

-- country_name    diff_forest_area
-- China   527229.062
-- United States   79200
-- India   69213.9844
-- Russian Federation  59395
-- Vietnam 55390

-- LARGEST CONCERNS
SELECT country_name,
       region,
       ROUND(CAST(forest_area_sqkm_2016 - forest_area_sqkm_1990 AS NUMERIC), 2)
diff_forest_area
FROM country_detail
```

```sql
WHERE country_name != 'World'
ORDER BY diff_forest_area ASC
LIMIT 5;

-- country_name      region diff_forest_area
-- 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


----------------------------------------------------------------
-- 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?

--- SUCCESS STORIES
SELECT country_name,
       ROUND(CAST((forest_area_sqkm_2016 - forest_area_sqkm_1990) /
forest_area_sqkm_1990 AS NUMERIC) * 100, 2) forest_area_pct_change
FROM country_detail
WHERE country_name != 'World'
ORDER BY forest_area_pct_change DESC
LIMIT 5;

-- country_name      forest_area_pct_change
-- Iceland 213.66
-- French Polynesia      181.82
-- Bahrain 177.27
-- Uruguay 134.11
-- Dominican Republic  82.46

-- LARGEST CONCERNS
SELECT country_name,
       region,
       ROUND(CAST((forest_area_sqkm_2016 - forest_area_sqkm_1990) /
forest_area_sqkm_1990 AS NUMERIC) * 100, 2) forest_area_pct_change
FROM country_detail
WHERE country_name != 'World'
ORDER BY forest_area_pct_change ASC
LIMIT 5;

-- country_name      region forest_area_pct_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


----------------------------------------------------------------
-- c. If countries were grouped by percent forestation in quartiles, which group had
the most countries in it in 2016?
WITH t_quartile AS (
    SELECT country_name, CASE WHEN pct_forest <= 25 THEN '0-25%'
                              WHEN pct_forest <= 50 AND pct_forest> 25 THEN '25-50%'
                              WHEN pct_forest <= 75 AND pct_forest> 50 THEN '50-75%'
                              ELSE '75-100%'
                         END quartile
    FROM forestation
    WHERE pct_forest IS NOT NULL AND country_name != 'World' AND year = 2016)
SELECT quartile, COUNT(country_name)
FROM t_quartile
```

```sql
GROUP BY quartile
ORDER BY quartile ASC;

-- quartile    count
-- 0-25%   85
-- 25-50%  72
-- 50-75%  38
-- 75-100% 9


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

SELECT country_name, ROUND(CAST(pct_forest AS NUMERIC), 2) pct_forest
FROM forestation
WHERE pct_forest > 75 AND country_name != 'World' AND year = 2016;

-- country_name    pct_forest
-- American Samoa  87.50
-- Micronesia, Fed. Sts.   91.86
-- Gabon   90.04
-- Guyana  83.90
-- Lao PDR 82.11
-- Palau   87.61
-- Solomon Islands 77.86
-- Suriname    98.26
-- Seychelles  88.41
-----------------------------------------------------------------
-- e. How many countries had a percent forestation higher than the United States in
2016?
SELECT COUNT(country_name) country_num
FROM forestation
WHERE pct_forest > (SELECT pct_forest FROM forestation
                    WHERE country_name = 'United States' AND year=2016)
    AND country_name != 'World'
    AND year = 2016;

-- country_num
-- 94
```