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 sqkm in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 39958245.9 sqkm, a loss of 1324449 sqkm, or 3.21%.

The forest area lost over this time period is slightly more than the entire land area of 1279999.9891 sqkm listed for the year 2016 (which is PERU).

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

The only regions of the world that decreased in percent forest area from 1990 to 2016 were Sub-Saharan Africa (dropped from 30.67% to 28.79%) and Latin America & Caribbean (51.03% to 46.16%). 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 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 79200 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 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
4th Quartile	9
3rd Quartile	38
2nd Quartile	72
1st Quartile	85

The largest number of countries in 2016 were found in the 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	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

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?*

The Global situation of deforestation is definitely alarming, however, on the bright side, bigger countries like China, USA have shown a great level of achievement by taking the measurements to prevent deforestation and eventually increasing the forest area. If we look at the data region-wise, out of 7 regions, only two regions show a reduction in the forest area. However, that reduction is too huge that it causes a percentage drop for the whole world forest area.

We should focus on the countries from the region- Sub Saharan Africa and Latin America & Caribbean as these two regions show the forest percentage reduction in majority of the countries than other regions. The countries lacking in forest percentage in these regions should follow the suit as China. Even the United States should follow the measurements that China took that showed such improvement in bringing up the forest area. All the countries should place some strict policies to be followed by the people. Some simple measures to take are- Planting more trees, Using fewer paper goods, Recycling more, Re-usable bags, Eating sustainably, and many more way to help the environment.

Appendix: SQL queries used

Steps to Complete

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).

```
DROP VIEW IF EXISTS country_view;

DROP VIEW IF EXISTS regions_new;

DROP VIEW IF EXISTS forestation;

CREATE VIEW forestation
AS
  (SELECT f.country_code                AS country_code,
          f.country_name                AS country_name,
          f.year                        AS YEAR,
          r.region,
          r.income_group,
          f.forest_area_sqkm,
          l.total_area_sq_mi            AS total_land_area_sq_mi,
          ( l.total_area_sq_mi * 2.59 ) AS total_land_area_sqkm,
          Round(( ( f.forest_area_sqkm / ( l.total_area_sq_mi * 2.59 ) ) * 100 )
                ::
                numeric, 2)             AS forest_percentage
   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);


SELECT *
FROM   forestation;
```

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.

```
SELECT forest_area_sqkm,
       YEAR,
       country_name
FROM   forestation
WHERE  country_name = 'World'
       AND YEAR = '1990';
```

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 forest_area_sqkm,
       year,
       country_name
FROM   forestation
WHERE  country_name = 'World'
       AND year = '2016';
```

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

AND

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

```
WITH tab_1990
     AS (SELECT forest_area_sqkm AS forest_area_1990,
                country_name,
                country_code
         FROM   forestation
         WHERE  country_name = 'World'
                AND YEAR = '1990'),


tab_2016
     AS (SELECT forest_area_sqkm AS forest_area_2016,
                country_name,
                country_code
         FROM   forestation
         WHERE  country_name = 'World'
                AND YEAR = '2016')


SELECT ( tab_1990.forest_area_1990 - tab_2016.forest_area_2016 ) AS
       forest_area_change,
       Round(( ( tab_1990.forest_area_1990 -
                       tab_2016.forest_area_2016 ) / tab_1990.forest_area_1990 *
               100 )
             :: numeric, 2)                                        AS
       forest_percent_change
FROM   tab_1990
       JOIN tab_2016
         ON tab_1990.country_name = tab_2016.country_name
```

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 *
FROM   forestation
WHERE  year = '2016'
       AND total_land_area_sqkm <= 1324449
ORDER  BY total_land_area_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).Based on the table you created, ...

```
CREATE VIEW regions_view
AS
  (SELECT region,
          year,
          Round(( ( SUM(forest_area_sqkm) / SUM(total_land_area_sqkm) ) * 100 )
                ::
                NUMERIC
          , 2) AS forest_percentage
   FROM   forestation
   WHERE  year = '2016'
           OR year = '1990'
   GROUP  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?

```
#for the world in the year 2016.

SELECT forest_percentage,
       year,
       region
FROM   regions_view
WHERE  region = 'World'
       AND year = '2016’;


## for the region in 2016.

SELECT forest_percentage,
       year,
       region
FROM   regions_view
WHERE  year = '2016'
ORDER  BY 1 DESC;


## We can either scroll down to see the lowest percentage or run the query without DESC to see the region at the top with the lowest forest percentage.
```

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 forest_percentage,
       year,
       region
FROM   regions_view
WHERE  region = 'World'
       AND year = '1990'
        OR region = 'World'
           AND year = '2016';
```

```sql
SELECT forest_percentage,
       year,
       region
FROM   regions_view
WHERE  year = '1990'
        OR year = '2016'
ORDER  BY 1 DESC ;
```

The above ^^ query shows filtering with the year 2016 as well as 1990 for the region- World. We could write the query filtering both at the same time or as single filter also as I showed in the a. Part for the year 2016 prior to this.

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

```sql
WITH forest_percentage_1990
     AS (SELECT forest_percentage AS f_p_1990,
                year,
                region
         FROM   regions_view
         WHERE  year = '1990'),


     forest_percentage_2016
     AS (SELECT forest_percentage AS f_p_2016,
                year,
                region
         FROM   regions_view
         WHERE  year = '2016')

SELECT prior.f_p_1990,
       recent.f_p_2016,
       ( prior.f_p_1990 - recent.f_p_2016 ) AS change_in_Forest,
       prior.year                           AS year,
       prior.region                         region
FROM   forest_percentage_1990 prior
       JOIN forest_percentage_2016 recent
         ON prior.region = recent.region
ORDER  BY 3 ;
```

3. COUNTRY-LEVEL DETAIL

SUCCESS STORIES

```
CREATE VIEW country_view
AS
  (SELECT country_name,
          year,
          region,
          forest_area_sqkm,
          Round(( ( SUM(forest_area_sqkm) / SUM(total_land_area_sqkm) ) * 100 )
                ::
                NUMERIC
          , 2) AS forest_percentage
   FROM   forestation
   WHERE  year = '2016'
           OR year = '1990'
   GROUP  BY 1,
             2,
             3,
             4);

WITH country_forest_percent_1990
     AS (SELECT forest_area_sqkm  AS cfa_sqkm_1990,
                forest_percentage AS cfp_1990,
                year,
                region,
                country_name
         FROM   country_view
         WHERE  year = '1990'),

     country_forest_percent_2016
     AS (SELECT forest_area_sqkm  AS cfa_sqkm_2016,
                forest_percentage AS cfp_2016,
                year,
                region,
                country_name
         FROM   country_view
         WHERE  year = '2016')
```

```sql
SELECT cfp_1990.cfa_sqkm_1990,
       cfp_2016.cfa_sqkm_2016,
       Round((( cfp_1990.cfa_sqkm_1990 -
       cfp_2016.cfa_sqkm_2016 )) :: NUMERIC, 2) AS
       change_cfa_sqkm,
       cfp_1990.cfp_1990,
       cfp_2016.cfp_2016,
       Round(( ( cfp_1990.cfa_sqkm_1990 - cfp_2016.cfa_sqkm_2016 ) /
               cfp_1990.cfa_sqkm_1990 *
                    100 ) :: NUMERIC, 2)
       AS change_cfp,
       cfp_1990.country_name
FROM   country_forest_percent_1990 AS cfp_1990
       join country_forest_percent_2016 cfp_2016
         ON cfp_1990.country_name = cfp_2016.country_name
WHERE  ( cfp_1990.cfp_1990 - cfp_2016.cfp_2016 ) IS NOT NULL
ORDER  BY change_cfa_sqkm;
```

When checking the percentage change in the forest area from 1990-2016, we will change the last code in the main query as " ORDER BY change_cfp ; "

We will now keep the WITH subquery country_forest_percent_1990 and country_forest_percent_2016 used for the Success Story, however we will change the main query we used in the success story for part (a) (b) now.

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?

```sql
SELECT cfp_1990.cfa_sqkm_1990,
       cfp_2016.cfa_sqkm_2016,
       Round((( cfp_1990.cfa_sqkm_1990 -
       cfp_2016.cfa_sqkm_2016 )) :: NUMERIC, 2) AS
```

```
       change_cfa_sqkm,
       cfp_1990.cfp_1990,
       cfp_2016.cfp_2016,
       Round(( ( cfp_1990.cfa_sqkm_1990 - cfp_2016.cfa_sqkm_2016 ) /
               cfp_1990.cfa_sqkm_1990 *
                     100 ) :: NUMERIC, 2)
       AS change_cfp,
       cfp_1990.country_name,
       cfp_1990.region
FROM   country_forest_percent_1990 AS cfp_1990
       join country_forest_percent_2016 cfp_2016
         ON cfp_1990.country_name = cfp_2016.country_name
WHERE  ( cfp_1990.cfp_1990 - cfp_2016.cfp_2016 ) IS NOT NULL
       AND cfp_1990.country_name NOT LIKE 'World'
ORDER  BY change_cfa_sqkm DESC ;
```

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?

```
SELECT cfp_1990.cfa_sqkm_1990,
       cfp_2016.cfa_sqkm_2016,
       Round((( cfp_1990.cfa_sqkm_1990 -
       cfp_2016.cfa_sqkm_2016 )) :: NUMERIC, 2) AS
       change_cfa_sqkm,
       cfp_1990.cfp_1990,
       cfp_2016.cfp_2016,
       Round(( ( cfp_1990.cfa_sqkm_1990 - cfp_2016.cfa_sqkm_2016 ) /
               cfp_1990.cfa_sqkm_1990 *
                     100 ) :: NUMERIC, 2)
       AS change_cfp,
       cfp_1990.country_name,
       cfp_1990.region
FROM   country_forest_percent_1990 AS cfp_1990
       join country_forest_percent_2016 cfp_2016
         ON cfp_1990.country_name = cfp_2016.country_name
WHERE  ( cfp_1990.cfp_1990 - cfp_2016.cfp_2016 ) IS NOT NULL
       AND cfp_1990.country_name NOT LIKE 'World'
ORDER  BY change_cfp DESC ;
```

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

```
WITH quartile
     AS (SELECT country_name,
                region,
                forest_percentage,
                forest_area_sqkm,
                CASE
                  WHEN forest_percentage < 25.00 THEN '1st Quartile'
                  WHEN forest_percentage < 50.00 THEN '2nd Quartile'
                  WHEN forest_percentage < 75.00 THEN '3rd Quartile'
                  ELSE '4th Quartile'
                END AS Quartiles
         FROM   forestation
         WHERE  forest_percentage IS NOT NULL
                AND country_name NOT LIKE 'World'
                AND year = '2016'
         ORDER  BY 3 DESC)
SELECT Count(*) nbr_of_countries,
       quartiles
FROM   quartile
GROUP  BY 2
ORDER  BY quartiles DESC
```

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

```
WITH quartile
     AS (SELECT country_name,
                forest_percentage,
                forest_area_sqkm,
                CASE
                  WHEN forest_percentage < 25.00 THEN '1st Quartile'
                  WHEN forest_percentage < 50.00 THEN '2nd Quartile'
                  WHEN forest_percentage < 75.00 THEN '3rd Quartile'
                  ELSE '4th Quartile'
                END AS Quartiles
         FROM   forestation
         WHERE  forest_percentage IS NOT NULL
                AND country_name NOT LIKE 'World'
                AND year = '2016')
```

```
SELECT *
FROM   quartile
ORDER  BY forest_percentage DESC
```

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

```
WITH quartile
     AS (SELECT country_name,
                region,
                forest_percentage,
                forest_area_sqkm,
                CASE
                  WHEN forest_percentage < 25.00 THEN '1st Quartile'
                  WHEN forest_percentage < 50.00 THEN '2nd Quartile'
                  WHEN forest_percentage < 75.00 THEN '3rd Quartile'
                  ELSE '4th Quartile'
                END AS Quartiles
         FROM   forestation
         WHERE  forest_percentage IS NOT NULL
                AND country_name NOT LIKE 'World'
                AND year = '2016')
SELECT Count(country_name)as nbr_of_countries
FROM   quartile
WHERE  forest_percentage > (SELECT forest_percentage
                           FROM   quartile
                           WHERE  country_name LIKE 'United States')
```