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 **3.31** %.

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
East Asia & Pacific	**25.78**	**26.36**
Europe & Central Asia	**37.28**	**38.04**
Latin America & Caribbean	**51.03**	**46.16**
Middle East & North Africa	**1.78**	**2.07**
North America	**35.65**	**36.04**
South Asia	**16.51**	**17.51**
Sub-Saharan Africa	**36.67**	**28.79**
World	**32.42**	**31.38**

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** (**36.67**% 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 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 **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 **344**% 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
0 - 25%	**84**
25 - 50%	**73**
50 - 75%	**38**
75 - 100%	**9**

The largest number of countries in 2016 were found in the 0-25% quartile.

There were **nine** 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?*
  *Overall, we could observe a increase in deforestation over the time period from 1990 to 2016 (32.42 and 31.38 respectively). However, it is interesting to note that most of the countries have taken actions against deforestation and increased in forest area. Surprisingly, **Iceland** increased in forest area by **344%** from 1990 to 2016. On the other hand, countries in **Sub-Saharan Africa** displayed an increase in deforestation compared to other regions. Moreover, it is observable that, yet by 2016 most of the countries in the world should take actions against deforestation. According to the Count of Countries Grouped by Forestation Percent Quartiles, 2016, most of the countries does not have at least 50% of the total land area allocated as forests.*

- *Which countries should we focus on over others?*
  Togo, Nigeria, Uganda, Mauritania, Honduras, Brazil, Indonesia, Myanmar, Tanzania

5. APPENDIX: SQL QUERIES USED

1. **Creating View as per the instructions**

```
CREATE VIEW forestation AS (

SELECT f.country_code AS country_code,

                    f.country_name AS country_name,

                   f.year AS year,

                    f.forest_area_sqkm AS forest_area_sqkm,

                    l.total_area_sq_mi * 2.59 AS total_land_area_sqkm,

                    l.total_area_sq_mi * 2.59 * 100 /f.forest_area_sqkm  AS
percent_of_total,

                    r.region AS region,
```

```
                    r.income_group AS income_group
FROM forest_area AS f
JOIN land_area AS l
ON f.country_code = l.country_code
        AND f.year = l.year
JOIN regions AS r
ON f.country_code = r.country_code)
;
```

(Using the query SELECT * FROM forestation LIMIT 5; , we can view the forestation relation)

2. **GLOBAL SITUATION**

Query to display forest area in the world in 1990.

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

Query to display forest area in the world in 1990.

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

Query to calculate the difference of forest area from 1990 to 2016 in the world

```
SELECT fn.country_name,
        fn.year,
        fn.forest_area_sqkm,
        fn.forest_area_sqkm - first_value(fn.forest_area_sqkm) OVER (ORDER BY
fn.year) AS diff_in_forest_area  – This will give us the difference
    FROM   forestation AS fn
    WHERE year IN (2016, 1990) – compare years
    AND country_name = 'World';
```

We used the same query above and changed it to find the percentage difference.

```
SELECT fn.country_name,
    fn.year,
    fn.forest_area_sqkm,
    fn.forest_area_sqkm - first_value(fn.forest_area_sqkm) OVER (ORDER BY fn.year)
AS diff_in_forest_area,
– This percentage value has not rounded off. If needed we can round off
    (fn.forest_area_sqkm - first_value(fn.forest_area_sqkm) OVER (ORDER BY
fn.year))*100/  fn.forest_area_sqkm AS percent_diff_forest_area
     FROM   forestation AS fn
     WHERE year IN (2016, 1990)
      AND country_name = 'World'
      ;
```

To round off, replace with :

Round(((fn.forest_area_sqkm - first_value(fn.forest_area_sqkm) OVER (ORDER BY fn.year))*100/ fn.forest_area_sqkm)::numeric, 2) AS percent_diff_forest_area

Query to find the country closest to total land area of 1350000sqkm. We just rounded up to the highest and check.

```
SELECT fn.country_name,
        fn.year,
          fn.total_land_area_sqkm
FROM   forestation AS fn
WHERE fn.year =2016
AND
fn.total_land_area_sqkm < 1350000
ORDER BY fn.total_land_area_sqkm DESC     — This will display the highest total land area less than 1350000sqkm
LIMIT 5;
```

2. REGIONAL OUTLOOK

Create CTE for 1990, 2016 and joined table for 1990 and 2016

```
WITH forest_percentage_1990 AS (
 SELECT
                fn.region,
                --fn.forest_area_sqkm,
                ROUND((SUM(fn.forest_area_sqkm) * 100 /
SUM(fn.total_land_area_sqkm))::numeric,2) AS percent_forest     – calculated the percentage of forest area out of the total land area
 FROM forestation AS fn
```

```
  WHERE fn.year = 1990
  GROUP BY fn.region
)
 ,
          forest_percentage_2016 AS (
  SELECT
                              fn.region,
                              --fn.forest_area_sqkm,
                              ROUND ((SUM(fn.forest_area_sqkm) * 100 /
SUM(fn.total_land_area_sqkm))::numeric,2) AS percent_forest
  FROM forestation AS fn
  WHERE fn.year = 2016
  GROUP BY  fn.region
)
,
   joinded_1990_2016 AS (
          SELECT
                               f90.region,
                               f90.percent_forest AS percent_forest_1990,
                               f16.percent_forest AS percent_forest_2016
```

```
        FROM forest_percentage_1990 AS f90
        JOIN forest_percentage_2016 AS f16
        ON f90.region = f16.region
)
```

Using SELECT to extract data from created CTEs grouped by Region with percentages of forest areas

```
SELECT
        region,
    jn.percent_forest_1990,
    jn.percent_forest_2016
FROM
        joinded_1990_2016 AS jn
GROUP BY
        region, jn.percent_forest_1990, jn.percent_forest_2016
ORDER BY
        region;
```

3. COUNTRY-LEVEL DETAIL

A. SUCCESS STORIES

We used the same query below (A2) with the following ORDER BY clause to display the amount difference of forest area between 1990 and 2016 and arranged it from highest to lowest to get the country with increased forest area. Query will display top 5.

```
ORDER BY
   jn.difference_total_forest_area DESC
LIMIT 5;
```

B. LARGEST CONCERNS

(A2) Display percent of change in forest area between 1990 and 2016 ordered from highest to lowest. (- shows decrease in forest area) Top 5 Amount Decrease in Forest Area by Country, 1990 & 2016:

```
WITH forest_area_1990 AS (
 SELECT
                fn.country_name,
                fn.forest_area_sqkm
 FROM forestation AS fn
 WHERE fn.year = 1990
 GROUP BY fn.country_name, fn.forest_area_sqkm
)
,
        forest_area_2016 AS (
 SELECT
                fn.country_name,
                fn.forest_area_sqkm
                FROM forestation AS fn
```

```
  WHERE fn.year = 2016

  GROUP BY  fn.country_name,fn.forest_area_sqkm

)

,

   joinded_forest_area_1990_2016 AS (

      SELECT

                  fa90.country_name,

                  fa90.forest_area_sqkm  AS total_forest_area_1990,

                  fa16.forest_area_sqkm  AS total_forest_area_2016,

                  fa16.forest_area_sqkm - fa90.forest_area_sqkm AS
difference_total_forest_area,

                  ROUND((((fa16.forest_area_sqkm -
fa90.forest_area_sqkm)*100)/fa90.forest_area_sqkm)::numeric,2) AS
percent_total_forest_area

            FROM forest_area_1990 AS fa90

            JOIN forest_area_2016 AS fa16

            ON fa90.country_name = fa16.country_name

)


SELECT

    country_name,
```

```
    jn.total_forest_area_1990,
    jn.total_forest_area_2016,
    jn.difference_total_forest_area,
    jn.percent_total_forest_area
FROM joinded_forest_area_1990_2016 AS jn
WHERE jn.percent_total_forest_area != 0
GROUP BY
    country_name,
    jn.total_forest_area_1990,
    jn.total_forest_area_2016,
    jn.difference_total_forest_area,
    jn.percent_total_forest_area
ORDER BY
    jn.percent_total_forest_area
LIMIT 5;
```

Complete query to display Top 5 (LIMIT =5) countries, ordered by the percent in change (1990 & 2016) in forest area

```
WITH forest_area_1990 AS (
  SELECT fn.region,
                fn.country_name,
                fn.forest_area_sqkm
```

```
 FROM forestation AS fn
 WHERE fn.year = 1990
 GROUP BY fn.region, fn.country_name, fn.forest_area_sqkm
)
 ,
          forest_area_2016 AS (
 SELECT fn.region,
                    fn.country_name,
                    fn.forest_area_sqkm
                    FROM forestation AS fn
 WHERE fn.year = 2016
 GROUP BY  fn.region, fn.country_name,fn.forest_area_sqkm
)
,
  joinded_forest_area_1990_2016 AS (
     SELECT fa90.region,
                    fa90.country_name,
                    fa90.forest_area_sqkm  AS total_forest_area_1990,
                    fa16.forest_area_sqkm  AS total_forest_area_2016,
```

```
                fa16.forest_area_sqkm - fa90.forest_area_sqkm AS
difference_total_forest_area,

                ROUND((((fa16.forest_area_sqkm -
fa90.forest_area_sqkm)*100)/fa90.forest_area_sqkm)::numeric,2) AS
percent_total_forest_area

        FROM forest_area_1990 AS fa90

        JOIN forest_area_2016 AS fa16

        ON fa90.country_name = fa16.country_name

)



SELECT

    country_name,

    region,

    jn.percent_total_forest_area

FROM joinded_forest_area_1990_2016 AS jn

WHERE jn.difference_total_forest_area != 0

GROUP BY

    country_name,

    jn.region,

    jn.total_forest_area_1990,

    jn.total_forest_area_2016,
```

```
    jn.difference_total_forest_area,
    jn.percent_total_forest_area
ORDER BY
    jn.percent_total_forest_area
LIMIT 5;
```

C. Quartiles

In order to calculate the figures for 2019, removed the CTE for 2019.

```
WITH forest_area_2016 AS (
  SELECT fn.region,
                    fn.country_name,
                    fn.forest_area_sqkm,
        fn.total_land_area_sqkm,
       ROUND((fn.forest_area_sqkm *100/fn.total_land_area_sqkm)::numeric,2) AS
percent_for_forests
                    FROM forestation AS fn
  WHERE fn.year = 2016
  GROUP BY  fn.country_name,fn.forest_area_sqkm, fn.region, fn.total_land_area_sqkm
)
 –Use case to group according to percentage
SELECT
```

```
(CASE WHEN fr.percent_for_forests BETWEEN 0 AND 25 THEN '25%'
        WHEN fr.percent_for_forests BETWEEN 25 AND 50 THEN '50%'
        WHEN fr.percent_for_forests BETWEEN 50 AND 75 THEN '75%'
        WHEN fr.percent_for_forests BETWEEN 75 AND 100 THEN '100%'
 ELSE 'NULL'
 END) AS quartile,
COUNT(*)    – Count the number of countries in each quartile
FROM forest_area_2016 as fr
WHERE percent_for_forests != 0   – Remove rows with missing values
GROUP BY quartile
ORDER BY quartile
;
```

We altered the same query, to display top quartile countries as below.

```
        WITH forest_area_2016 AS (
          SELECT fn.region,
                        fn.country_name,
                        fn.forest_area_sqkm,
              fn.total_land_area_sqkm,
              ROUND((fn.forest_area_sqkm *100/fn.total_land_area_sqkm)::numeric,2)
        AS percent_for_forests
```

```
                    FROM forestation AS fn
  WHERE fn.year = 2016
  GROUP BY  fn.country_name,fn.forest_area_sqkm, fn.region,
fn.total_land_area_sqkm
)
SELECT
(CASE WHEN fr.percent_for_forests BETWEEN 0 AND 25 THEN '25%'
      WHEN fr.percent_for_forests BETWEEN 25 AND 50 THEN '50%'
      WHEN fr.percent_for_forests BETWEEN 50 AND 75 THEN '75%'
      WHEN fr.percent_for_forests BETWEEN 75 AND 100 THEN '100%'
 ELSE 'NULL'
 END) AS quartile,
 country_name,
 region,
 percent_for_forests
FROM forest_area_2016 as fr
WHERE percent_for_forests != 0 – For a better view removed missing values
ORDER BY percent_for_forests DESC
LIMIT 9;  – since we have 9 countries in the top quartile
```