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 **41282694.90 sqKm** in 1990. As of 2016, the most recent year for which data was available, that number had fallen to **39958245.90 sqKm**, a loss of **1324449.00 sqKm** , or **3.2 % decreses.**

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 **1279999.99 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
Latin America & Caribbean	51.03	46.16
Europe & Central Asia	37.27	38.06
North America	36.65	36.04
World	32.42	31.38
Sub-Saharan Africa	32.19	27.56
East Asia & Pacific	25.77	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 **Sub-Saharan Africa** (dropped from **32.19 % to 27.56 %**) 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 **527229.06 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**, but it only saw an increase of **79200.00 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 **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	59.27
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
FIRST (0-25 %)	85
Second (25-50 %)	73
Third (50-75 %)	38
Fourth (75-100 %)	9

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

According to the world bank there is lot decrease in forest area from 1990 to 2016 for some countries and regions. The highest relative forestation was in Latin America & Caribbean region in 1990 as well as 2016. Togo, Nigeria, Uganda, and Mauritania are the countries from region Sub-saharan Africa where forest area decreased. When we look at count of countries grouped by forestation percentage in 1990 and 2016, we can see that there are 85 countries falls in first quartile between 0-25%.

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

The focus should be given in the region Sub-Saharan Africa where 4 out 5 countries are in top 5 percent decrease in forest area, the countries are Togo, Nigeria, Uganda, and Mauritania. We need to study about these countries and should identify the reasons of forestation. China is a country where forest area was increased by 527229.06 sqkm from 1990 to 2016. So, we need to study about what changes are done in this country over time.

APPENDIX: SQL queries used

1. Create view and create column per_forest_land

```
DROP VIEW IF EXISTS 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,
  ROUND((SUM(f.forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)):: NUMERIC,2) AS
per_forest_land
FROM forest_area f
INNER JOIN land_area l
ON f.country_code = l.country_code AND f.year = l.year
INNER JOIN regions r
ON f.country_code = r.country_code
GROUP BY f.country_code, f.country_name, f.year, f.forest_area_sqkm,
l.total_area_sq_mi,r.region, r.income_group;
```

1. GLOBAL SITUATION

1. **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 ROUND(SUM(forest_area_sqkm):: NUMERIC,2)
FROM forestation
WHERE year = 1990 AND region = 'World'
```

2. **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 denoted as “World" in the region table.**

```
SELECT ROUND(SUM(forest_area_sqkm ):: NUMERIC,2)
FROM forestation
WHERE year = 2016 AND region = 'World'
```

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

```
WITH T1 AS
(SELECT ROUND(SUM(forest_area_sqkm):: NUMERIC,2) AS sum_1990
FROM forestation
WHERE year = 1990 AND region = 'World'),
T2 AS
(SELECT ROUND(SUM(forest_area_sqkm):: NUMERIC,2) AS sum_2016
FROM forestation
WHERE year = 2016 AND region = 'World')

SELECT sum_1990 - sum_2016 AS change_forest_area
FROM T1,T2
```

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

```
WITH T1 AS
(SELECT ROUND(SUM(forest_area_sqkm):: NUMERIC,2) AS sum_1990
FROM forestation
WHERE year = 1990 AND region = 'World'),
T2 AS
(SELECT ROUND(SUM(forest_area_sqkm):: NUMERIC,2) AS sum_2016
FROM forestation
WHERE year = 2016 AND region = 'World')

SELECT ROUND((sum_1990 - sum_2016) / sum_1990 :: NUMERIC,3)  * 100 AS
percent_change_forest_area
FROM T1,T2
GROUP BY sum_1990, sum_2016
```

5. **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, total_area_sq_mi,
        ROUND((total_area_sq_mi * 2.59) :: NUMERIC,2) total_land_area
FROM forestation
WHERE year = 2016 AND total_area_sq_mi IS NOT NULL AND (total_area_sq_mi * 2.59) <=
1324449
GROUP BY 1,2
ORDER BY 3 DESC
LIMIT 1
```

2. REGIONAL OUTLOOK

1. **A) What was the percent forest of the entire world in 2016?**

```
SELECT region, ROUND((SUM(forest_area_sqkm * 100) / SUM(total_area_sq_mi *
2.59))::NUMERIC, 2) AS forest_percent
 FROM forestation
 WHERE year = 2016 AND region = 'World' AND forest_area_sqkm IS NOT NULL AND
total_area_sq_mi IS NOT NULL
```

1. B) 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, ROUND((SUM(forest_area_sqkm * 100) / SUM(total_area_sq_mi *
2.59))::NUMERIC, 2) AS forest_percent
 FROM forestation
 WHERE year = 2016 AND forest_area_sqkm IS NOT NULL AND total_area_sq_mi IS NOT NULL
 GROUP BY 1
 ORDER BY 2 DESC
```

2. **A) What was the percent forest of the entire world in 1990?**

```
SELECT region, ROUND((SUM(forest_area_sqkm * 100) / SUM(total_area_sq_mi *
2.59))::NUMERIC, 2) AS forest_percent
 FROM forestation
 WHERE year = 1990 AND region = 'World' AND forest_area_sqkm IS NOT NULL AND
total_area_sq_mi IS NOT NULL
```

2. 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, ROUND((SUM(forest_area_sqkm * 100) / SUM(total_area_sq_mi *
2.59))::NUMERIC, 2) AS forest_percent
 FROM forestation
 WHERE year = 1990 AND forest_area_sqkm IS NOT NULL AND total_area_sq_mi IS NOT NULL
 GROUP BY 1
 ORDER BY 2 DESC
```

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

```
WITH T1 AS
(SELECT region, ROUND((SUM(forest_area_sqkm * 100) / SUM(total_area_sq_mi *
2.59))::NUMERIC, 2) AS forest_percent_1990
 FROM forestation
 WHERE year = 1990 AND forest_area_sqkm IS NOT NULL AND total_area_sq_mi IS NOT NULL
 GROUP BY 1),

 T2 AS
 (SELECT region, ROUND((SUM(forest_area_sqkm * 100) / SUM(total_area_sq_mi *
2.59))::NUMERIC, 2) AS forest_percent_2016
 FROM forestation
 WHERE year = 2016 AND forest_area_sqkm IS NOT NULL AND total_area_sq_mi IS NOT NULL
 GROUP BY 1)

 SELECT T1.region, forest_percent_1990, forest_percent_2016, ROUND(((forest_percent_1990
- forest_percent_2016) * 100/ forest_percent_1990)::NUMERIC, 2) AS percent_decrease
 FROM T1
 JOIN T2
 ON T1.region = T2.region
 GROUP BY 1,2,3
 ORDER BY 4  DESC
```

3. COUNTRY-LEVEL DETAIL

1. **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_name, forest_area_sqkm forest_area_1990
 FROM forestation
 WHERE year = 1990 AND forest_area_sqkm IS NOT NULL
```

```
),
T2 AS
(SELECT country_name, forest_area_sqkm  forest_area_2016
 FROM forestation
 WHERE year = 2016 AND forest_area_sqkm IS NOT NULL
 )
 SELECT T1.country_name,forest_area_1990, forest_area_2016, ROUND((forest_area_2016 -
forest_area_1990):: NUMERIC,2) AS forest_area_increase
 FROM T1
 JOIN T2
 ON T1.country_name = T2.country_name
GROUP BY 1,2,3
ORDER BY 4 DESC
LIMIT 5
```

2. **A) 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 T1 AS
(SELECT country_name, (SUM(forest_area_sqkm) * 100/SUM(total_area_sq_mi * 2.59))
percent_forest_area_1990
 FROM forestation
 WHERE year = 1990 AND forest_area_sqkm IS NOT NULL AND total_area_sq_mi IS NOT NULL
 GROUP BY 1
 ),
T2 AS
(SELECT country_name, (SUM(forest_area_sqkm) * 100/SUM(total_area_sq_mi * 2.59))
percent_forest_area_2016
 FROM forestation
 WHERE year = 2016 AND forest_area_sqkm IS NOT NULL AND total_area_sq_mi IS NOT NULL
 GROUP BY 1
 )
 SELECT T1.country_name, percent_forest_area_1990, percent_forest_area_2016,
ROUND(((percent_forest_area_1990 -
percent_forest_area_2016)*100/percent_forest_area_1990):: NUMERIC,2) AS
percent_forest_area_decrease
 FROM T1
 JOIN T2
 ON T1.country_name = T2.country_name
ORDER BY 4
LIMIT 5
```

Table 3.1
-Top 5 Amount Decrease in Forest Area by Country, 1990 & 2016:

```
WITH T1 AS
(SELECT country_name, region, forest_area_sqkm forest_area_1990
 FROM forestation
```

```
 WHERE year = 1990 AND forest_area_sqkm IS NOT NULL
 ),
T2 AS
(SELECT country_name, region, forest_area_sqkm  forest_area_2016
 FROM forestation
 WHERE year = 2016 AND forest_area_sqkm IS NOT NULL
 )
 SELECT T1.country_name, T1.region, forest_area_1990, forest_area_2016,
ROUND((forest_area_1990 - forest_area_2016):: NUMERIC,2) AS forest_area_change
 FROM T1
 JOIN T2
 ON T1.country_name = T2.country_name
 WHERE T1.country_name NOT LIKE 'World'
GROUP BY 1,2,3,4
ORDER BY 5 DESC
LIMIT 5
```

Table 3.2
Top 5 Percent Decrease in Forest Area by Country, 1990 & 2016

```
WITH T1 AS
(SELECT country_name, (SUM(forest_area_sqkm) * 100/SUM(total_area_sq_mi * 2.59))
percent_forest_area_1990
 FROM forestation
 WHERE year = 1990 AND forest_area_sqkm IS NOT NULL AND total_area_sq_mi IS NOT NULL
 GROUP BY 1
 ),
T2 AS
(SELECT country_name, (SUM(forest_area_sqkm) * 100/SUM(total_area_sq_mi * 2.59))
percent_forest_area_2016
 FROM forestation
 WHERE year = 2016 AND forest_area_sqkm IS NOT NULL AND total_area_sq_mi IS NOT NULL
 GROUP BY 1
 )
 SELECT T1.country_name, percent_forest_area_1990, percent_forest_area_2016,
ROUND(((percent_forest_area_1990 -
percent_forest_area_2016)*100/percent_forest_area_1990):: NUMERIC,2) AS
percent_forest_area_decrease
 FROM T1
 JOIN T2
 ON T1.country_name = T2.country_name
ORDER BY 4 DESC
LIMIT 5
```

C : Quartiles

Table 3.3

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

```
WITH T1 AS
(SELECT country_name, ROUND((SUM(forest_area_sqkm) * 100/SUM(total_area_sq_mi *
2.59)) :: NUMERIC,2) percent_forest_area_2016
 FROM forestation
 WHERE year = 2016 AND forest_area_sqkm IS NOT NULL AND total_area_sq_mi IS NOT NULL
 GROUP BY 1
 ),
 T2 AS
(SELECT T1.country_name,
            CASE WHEN percent_forest_area_2016 > 75 THEN 'Fourth'
                        WHEN percent_forest_area_2016 <=75 AND percent_forest_area_2016 > 50
THEN 'Third'
                        WHEN percent_forest_area_2016 < 50 AND percent_forest_area_2016 > 25
THEN 'Second'
                        ELSE 'First'
            END AS quartiles
  FROM T1
)

SELECT   quartiles quratile_group, COUNT(*) AS city_count
FROM T2
GROUP BY 1
ORDER BY 2 DESC
```

Table 3.4

-- Top Quartile Countries, 2016. List of countries and their respective forest land, denoted as a percentage .

```
WITH T1 AS
(SELECT country_name, region, ROUND((SUM(forest_area_sqkm) *
100/SUM(total_area_sq_mi * 2.59)) :: NUMERIC,2) percent_forest_area_2016
 FROM forestation
 WHERE year = 2016 AND forest_area_sqkm IS NOT NULL AND total_area_sq_mi IS
NOT NULL
 GROUP BY 1,2
 ),
 T2 AS
(SELECT T1.country_name, T1.region, percent_forest_area_2016,
            CASE WHEN percent_forest_area_2016 > 75 THEN 'Fourth'
```

```
                    WHEN percent_forest_area_2016 <=75 AND
percent_forest_area_2016 > 50 THEN 'Third'
                    WHEN percent_forest_area_2016 < 50 AND percent_forest_area_2016
> 25 THEN 'Second'
                    ELSE 'First'
            END AS quartiles
  FROM T1
)

SELECT   country_name, region, percent_forest_area_2016
FROM T2
WHERE quartiles = 'Fourth'
ORDER BY 3 DESC
```