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 **39958245.5 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 **Peru** listed for the year 2016 (which is **1279999.99 sqkm**).

2. **REGIONAL OUTLOOK**

In 2016, the percentage 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 the **Middle East & North Africa**, with **2.07%** forestation.

In 1990, the percentage of the total land area of the world designated as a 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 the **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**
Sub-Saharan Africa	**30.67**	**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** (**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 of forest area in 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 km^2**. 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 km^2**, 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	541,500.00 km^2
Indonesia	East Asia & Pacific	282,193.98 km^2
Myanmar	East Asia & Pacific	107,234.00 km^2
Nigeria	Sub-Saharan Africa	106,506.00 km^2
Tanzania	Sub-Saharan Africa	102,320.00 km^2

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%	85
25%-50%	72
50%-75%	38
75%-100%	9

The largest number of countries in 2016 were found in the **0-25(1st)** quartile.

There were **85** 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%

4. RECOMMENDATIONS

Write out a set of recommendations as an analyst on the ForestQuery team.

- *What have you learned from the World Bank data?*

    1. My first observation is a huge loss in the percentage of forest area. Within the span of 26 years (1990-2016), there is a forest area change of 3.2% which is a serious concern. This percentage lost is almost equal to the land area of Peru (2016).

2. The countries which decreased in forest area from 1990 to 2016 are Latin America & Caribbean (drops from 51.03% to 46.16%), Sub-Saharan Africa (drops from 30.67% to 28.79%) and the entire World has also a slight decrease in the forest area.
3. Countries like China, the United States, and India have increased in forest area. Even a small country 'Iceland' has a big percentage of increased forest area 213.66%.
4. In Table 3.1, we can observe the top 5 countries which are dramatically decreasing in their forest areas. Among them, Latin America & Caribbean ranks at the top which is a huge decrease of 541,500.00 km^2 in their land area.

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

1. Primary focus must be on Nigeria because it is the only country that ranks in the top 5 both in terms of absolute square kilometer decrease in the forest as well as the percent decrease in forest area from 1990 to 2016.
2. Countries like Brazil, Myanmar, Indonesia, and Tanzania also should be concerned because there is a huge decrease in their forest areas.
3. From the percentage of forest area point of view, Sub-Saharan Africa has to focus more to increase its forest area.
4. Latin America & Caribbean region has a huge percent drop of almost 5% from (1990-2016). This region also should be focused on.

5. APPENDIX: SQL Queries Used

1. CREATE VIEW OF "forestation" and other steps before starting the project:

```
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*2.59) AS total_area_sqkm,
r.region, r.income_group,
((SUM(forest_area_sqkm)/SUM(total_area_sq_mi*2.59))*100) AS percent_forest
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 r.country_code = l.country_code
GROUP BY 1,2,3,4,5,6,7);

SELECT *
FROM forestation;
```

PART 1: GLOBAL SITUATION

1. 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.

```
WITH total_forest_area_1990 AS
(SELECT SUM(forest_area_sqkm) AS total_forest_area_sqkm
FROM forest_area
WHERE year = 1990 AND country_name = 'World')

SELECT * FROM total_forest_Area_1990;
```

---

1. 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."

```
WITH total_forest_area_2016 AS
(SELECT SUM(forest_area_sqkm) AS total_forest_area_sqkm
FROM forest_area
WHERE year = 2016 AND country_name = 'World')

SELECT * FROM total_forest_Area_2016;
```

---

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

```
SELECT
(SELECT SUM(forest_area_sqkm) AS total_forest_area_sqkm
FROM forest_area
WHERE year = 1990 AND country_name = 'World') -
(SELECT SUM(forest_area_sqkm) AS total_forest_area_sqkm
FROM forest_area
WHERE year = 2016 AND country_name = 'World') AS change_in_forest_area
FROM forestation
LIMIT 1;
```

---

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

```
SELECT
(((SELECT SUM(forest_area_sqkm) AS total_forest_area_sqkm
FROM forest_area
WHERE year = 1990 AND country_name = 'World') -
(SELECT SUM(forest_area_sqkm) AS total_forest_area_sqkm
FROM forest_area
WHERE year = 2016 AND country_name = 'World')) /
(SELECT SUM(forest_area_sqkm) AS total_forest_area_sqkm
FROM forest_area
WHERE year = 1990 AND country_name = 'World'))*100 AS percent_change_in_forest_area
FROM forestation
LIMIT 1;
```

1. 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 country_name, total_area_sqkm,
ABS((total_area_sqkm - (SELECT
(SELECT SUM(forest_area_sqkm) AS total_forest_area_sqkm
FROM forest_area
WHERE year = 1990 AND country_name = 'World') -
(SELECT SUM(forest_area_sqkm) AS total_forest_area_sqkm
FROM forest_area
WHERE year = 2016 AND country_name = 'World'))))) AS forest_area_lost
FROM forestation
WHERE year = 2016
GROUP BY 1,2
ORDER BY 3
LIMIT 1;
```

PART 2: Regional Outlook

1. CREATE TABLE that shows regions and their percent forest area in 1990 and 2016.

YEAR = 2016

```
WITH forest_percentage_2016
AS
(SELECT region,
ROUND(((SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100)::NUMERIC, 2) AS
percent_forest_area
FROM forestation
WHERE year = 2016
GROUP BY 1
ORDER BY 2 DESC)

SELECT * FROM forest_percentage_2016;
```

a. 1 What was the percent forest of the entire world in 2016?

```
SELECT * FROM forest_percentage_2016
WHERE region = 'World';
```

a. 2 Which region had the HIGHEST percent forest in 2016?

```
SELECT * FROM forest_percentage_2016
ORDER BY 2 DESC
LIMIT 1;
```

a.3 Which region had the LOWEST percent forest in 2016?

```
SELECT * FROM forest_percentage_2016
ORDER BY 2
LIMIT 1;
```

YEAR = 1990

```
WITH forest_percentage_1990
AS
(SELECT region,
ROUND(((SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100)::NUMERIC, 2) AS
percent_forest_area
FROM forestation
WHERE year = 1990
GROUP BY 1
ORDER BY 2 DESC)

SELECT * FROM forest_percentage_1990;
```

b.1 What was the percent forest of the entire world in 1990?

```
SELECT * FROM forest_percentage_1990
WHERE region = 'World';
```

b.2 Which region had the HIGHEST percent forest in 1990?

```
SELECT * FROM forest_percentage_1990
ORDER BY 2 DESC
LIMIT 1;
```

b.3 Which region had the LOWEST percent forest in 1990?

```
SELECT * FROM forest_percentage_1990
ORDER BY 2
LIMIT 1;
```

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

```
WITH forest_percentage_2016
AS
(SELECT region,
ROUND(((SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100)::NUMERIC, 2) AS
percent_forest_area
FROM forestation
WHERE year = 2016
```

```
GROUP BY 1),
forest_percentage_1990
AS
(SELECT region,
ROUND(((SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100)::NUMERIC, 2) AS
percent_forest_area
FROM forestation
WHERE year = 1990
GROUP BY 1)

SELECT fp1990.region,
ROUND(fp1990.percent_forest_area::NUMERIC, 2) AS percent_fa_1990,
ROUND(fp2016.percent_forest_area::NUMERIC, 2) AS percent_fa_2016
FROM forest_percentage_1990 AS fp1990
JOIN forest_percentage_2016 AS fp2016
ON fp1990.region = fp2016.region
WHERE fp1990.percent_forest_area > fp2016.percent_forest_area
GROUP BY 1,2,3;
```

PART 3: Country-Level Detail

a. Which 5 countries saw the largest amount of decrease in forest area from 1990 to 2016? What was the difference in forest area for each?

```
WITH forest_change_1990
AS
(SELECT region, country_name, SUM(forest_area_sqkm) AS total_forest_area1
FROM forestation
WHERE year = 1990
GROUP BY 1,2),

forest_change_2016
AS
(SELECT region, country_name, SUM(forest_area_sqkm) AS total_forest_area2
FROM forestation
WHERE year = 2016
GROUP BY 1,2)

SELECT f1.country_name, f1.region, ROUND((f1.total_forest_area1-
f2.total_forest_area2)::NUMERIC,2) AS forest_change_sqkm
FROM forest_change_1990 f1
JOIN forest_change_2016 f2                                    ON f1.country_name =
f2.country_name
AND (f1.total_forest_area1 IS NOT NULL AND f2.total_forest_area2 IS NOT NULL) AND
f1.region != 'World'
ORDER BY 3 DESC
LIMIT 5;
```

---

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?

```
WITH forest_change_1990
AS
(SELECT region, country_name, SUM(forest_area_sqkm) AS total_forest_area1
FROM forestation
WHERE year = 1990
GROUP BY 1,2),

forest_change_2016
AS
(SELECT region, country_name, SUM(forest_area_sqkm) AS total_forest_area2
FROM forestation
WHERE year = 2016
GROUP BY 1,2)

SELECT f1.country_name, f1.region, ROUND((((f1.total_forest_area1-
f2.total_forest_area2)/(f1.total_forest_area1))*100)::NUMERIC,2) AS percent_forest_change
FROM forest_change_1990 f1
JOIN forest_change_2016 f2                                          ON f1.country_name =
f2.country_name
AND (f1.total_forest_area1 IS NOT NULL AND f2.total_forest_area2 IS NOT NULL) AND
f1.region != 'World'
ORDER BY 3 DESC
LIMIT 5;
```

---

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

```
WITH fc2016
AS
(SELECT region, country_name,
ROUND(((SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100)::NUMERIC, 2) AS
percent_forestation
FROM forestation
WHERE year = 2016 AND region != 'World' AND forest_area_sqkm IS NOT NULL AND
total_area_sqkm IS NOT NULL
GROUP BY 1,2),
quartiles
AS
(SELECT fc2016.region, fc2016.country_name,
 CASE WHEN fc2016.percent_forestation >= 75 THEN '75%-100%'
         WHEN fc2016.percent_forestation BETWEEN 50 AND 75 THEN '50%-75%'
      WHEN fc2016.percent_forestation BETWEEN 25 AND 50 THEN '25%-75%'
      WHEN fc2016.percent_forestation <= 25 THEN '0-25%'
         END AS percentile
 FROM fc2016
 ORDER BY 3 DESC)

 SELECT quartiles.percentile AS Quartiles,
             COUNT(quartiles.percentile) AS number_of_countries
```

```
FROM quartiles
GROUP BY 1
ORDER BY 2 DESC;
```

---

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

```
WITH fc2016
AS
(SELECT region, country_name,
ROUND(((SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100)::NUMERIC, 2) AS
percent_forestation
FROM forestation
WHERE year = 2016 AND region != 'World' AND forest_area_sqkm IS NOT NULL AND
total_area_sqkm IS NOT NULL
GROUP BY 1,2),
quartiles
AS
(SELECT fc2016.region, fc2016.country_name,
 CASE WHEN fc2016.percent_forestation >= 75 THEN '75%-100%'
        WHEN fc2016.percent_forestation BETWEEN 50 AND 75 THEN '50%-75%'
      WHEN fc2016.percent_forestation BETWEEN 25 AND 50 THEN '25%-75%'
      WHEN fc2016.percent_forestation <= 25 THEN '0-25%'
        END AS percentile
 FROM fc2016
 ORDER BY 3 DESC)

SELECT quartiles.country_name, quartiles.region,
ROUND((fc2016.percent_forestation)::NUMERIC, 2) AS Pct_Designated_Forest
FROM quartiles
JOIN fc2016
ON fc2016.country_name = quartiles.country_name
WHERE quartiles.percentile = '75%-100%'
GROUP BY 1,2,3
ORDER BY 1;
```

---

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

```
WITH fc2016
AS
(SELECT region, country_name,
ROUND(((SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100)::NUMERIC, 2) AS
percent_forestation
FROM forestation
WHERE year = 2016 AND region != 'World' AND forest_area_sqkm IS NOT NULL AND
total_area_sqkm IS NOT NULL
GROUP BY 1,2)

SELECT COUNT(fc2016.country_name) AS no_countries
FROM fc2016
WHERE fc2016.percent_forestation > (SELECT fc2016.percent_forestation FROM fc2016
WHERE country_name = 'United States');
```