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 41,282,694.9 km2 in 1990.As of 2016, the most recent year for which data was available, that number had fallen to 39,958,245.9 km2, a loss of 1,324,449 km2, or 3.21% decrease.

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 1,279,999.98 square kilometers).

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 and the Caribbean, with 46.16%, and the region with the lowest relative forestation was the Middle East and 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 and the Caribbean, with 51.03%, and the region with the lowest relative forestation was Middle East and 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 and 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 the 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 527,229.06 km2. 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 79,200 km2, much lower than the figure for China.

China and the U.S 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 and Caribbean	541,510 km2
Indonesia	East Asia and Pacific	282,193.98 km2
Myanmar	East Asia and Pacific	107,234 km2
Nigeria	Sub-Saharan Africa	106506 km2
Tanzania	Sub-Saharan Africa	102,320 km2

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	61.8%
Uganda	Sub-Saharan Africa	59.27%
Mauritania	Sub-Saharan Africa	46.75%
Honduras	Latin America and 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 and 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 mile 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%	73
50%-75%	38
75%-100%	9

The largest number of countries in 2016 were found in the bottom or (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
American Samoa	East Asia & Pacific	87.50
Gabon	Sub-Saharan Africa	90.04
Guyana	Latin America & Caribbean	83.90
Lao PDR	East Asia & Pacific	82.11
Micronesia, Fed. Sts.	East Asia & Pacific	91.86
Palau	East Asia & Pacific	87.61
Seychelles	Sub-Saharan Africa	88.41
Solomon Islands	East Asia & Pacific	77.86

Suriname	Latin America & Caribbean	98.26

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?*

I learned that many regions in the world are increasing in forest area. For example, Europe and Central Asia, North America, East Asia and the Pacific, South Asia, and Middle East and North Africa are all increasing in forest area and forest area percentage. However, that the world has decreased in forest area from 1990 to 2016 underscores how much forest Latin America and the Caribbean, and Sub-Saharan Africa have lost. I would recommend focusing on the countries with the largest absolute forest area change, found in table 3.1. The top 5 countries in this category are Brazil, Indonesia, Myanmar, Nigeria, and Tanzania. While looking at countries that have high percentage forestation decrease is useful, I think focusing on the countries that are reducing the most forest by land area is most important. Lastly, I think partnering with China to find best practices for increasing forest area would be very helpful.

5. APPENDIX: SQL Queries Used

--- 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 forest_area.country_code ,

        forest_area.year,

         forest_area.forest_area_sqkm,

         land_area.country_name, land_area.total_area_sq_mi, regions.region,
regions.income_group, forest_area.forest_area_sqkm / (land_area.total_area_sq_mi * 2.59) *
100 forest_percentage , land_area.total_area_sq_mi * 2.59 total_area_sqkm

 FROM forest_area

 JOIN land_area

 ON forest_area.country_code = land_area.country_code AND forest_area.year =
land_area.year

 JOIN regions

 ON forest_area.country_code = land_area.country_code);
```

--- 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.---

```
SELECT f.forest_area_sqkm, f.year
                    FROM forest_area f
       WHERE f.country_name = 'World'
          AND f.year = 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.”---

```
SELECT f.forest_area_sqkm, f.year
                    FROM forest_area f
       WHERE f.country_name = 'World'
          AND f.year = 2016;
```

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

```
SELECT sub1.forest_area_sqkm - sub2.forest_area_sqkm AS diff_forest_area_sq_km
     FROM (SELECT f.country_code AS cc, f.forest_area_sqkm
```

```
            FROM forest_area f
          WHERE f.country_name = 'World'
                AND f.year = 1990) AS sub1
    JOIN (SELECT f.country_code AS cc,f.forest_area_sqkm
                FROM forest_area f
          WHERE f.country_name = 'World'
                AND f.year = 2016) AS sub2
    ON sub1.cc = sub2.cc;
```

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

```
WITH area_2016
AS
(SELECT forest_area_sqkm AS a_2016, year
 FROM forestation
 WHERE country_name = 'World'
 AND year = 2016),
 area_1990
AS
```

```
(SELECT forest_area_sqkm AS a_1990, year

 FROM forestation

 WHERE country_name = 'World'

 AND year = 1990),

 diffs

 AS

 (SELECT a_2016, a_1990, a_2016 - a_1990 diff, (a_2016 - a_1990) / a_1990 * 100 AS
diff_percentage

  FROM area_2016, area_1990)

SELECT a_2016, a_1990, diff, ROUND(diff_percentage::NUMERIC,2) AS diff_percentage

FROM diffs;
```

---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 DISTINCT country_name, total_area_sqkm

FROM forestation

WHERE total_area_sqkm BETWEEN 1270000 AND 1350000;
```

---2. REGIONAL OUTLOOK---

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

```
SELECT country_name,
 Round(((Sum(forest_area_sqkm) / Sum(total_area_sq_mi*2.59))*100)::Numeric, 2) AS
percent_forest
FROM forestation
WHERE year = 2016
 AND country_name = 'World'
GROUP BY country_name;
```

---Which region had the HIGHEST percent forest in 2016, and which had the LOWEST, to 2 decimal places?---

------lowest----

```
SELECT region,
 Round(((Sum(forest_area_sqkm) / Sum(total_area_sq_mi*2.59))*100)::Numeric, 2) AS
percent_forest_2016
FROM forestation
WHERE year = 2016
GROUP BY region
```

```
ORDER BY percent_forest_2016;
```

---- highest----

```
SELECT region,
 Round(((Sum(forest_area_sqkm) / Sum(total_area_sq_mi*2.59))*100)::Numeric, 2) AS
percent_forest_2016
FROM forestation
WHERE year = 2016
GROUP BY region
ORDER BY percent_forest_2016 DESC;
```

---a.What was the percent forest of the entire world in 1990? ---

```
SELECT country_name,
 Round(((Sum(forest_area_sqkm) / Sum(total_area_sq_mi*2.59))*100)::Numeric, 2) AS
percent_forest
FROM forestation
WHERE year = 1990
 AND country_name = 'World'
```

```
GROUP BY country_name;
```

---Which region had the HIGHEST percent forest in 1990, and which had the LOWEST, to 2 decimal places?---

```
------lowest----

SELECT region,

 Round(((Sum(forest_area_sqkm) / Sum(total_area_sq_mi*2.59))*100)::Numeric, 2) AS

percent_forest_1990

FROM forestation

WHERE year = 1990

GROUP BY region

ORDER BY percent_forest_1990;

---- highest----

SELECT region,

 Round(((Sum(forest_area_sqkm) / Sum(total_area_sq_mi*2.59))*100)::Numeric, 2) AS

percent_forest_1990

FROM forestation

WHERE year = 1990

GROUP BY region
```

```
ORDER BY percent_forest_1990 DESC;
```

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

(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.)---

```
CREATE OR REPLACE VIEW regional_distr

AS

SELECT r.region,
```

```
l.year,
SUM(f.forest_area_sqkm) total_forest_area_sqkm,
SUM(l.total_area_sq_mi*2.59) AS total_area_sqkm,
 (SUM(f.forest_area_sqkm)/SUM(l.total_area_sq_mi*2.59))*100 AS percent_fa_region
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
GROUP BY 1,2
ORDER BY 1,2;
```

-----2.c. Based on the table you created, which regions of the world DECREASED in forest area from 1990 to 2016?----

```
WITH table1990 AS
(SELECT * FROM regional_distr
WHERE year =1990),
        table2016 AS
 (SELECT * FROM regional_distr WHERE year = 2016)
```

```
SELECT table1990.region,
       ROUND(CAST(table1990.percent_fa_region AS NUMERIC),2) AS fa_1990,
       ROUND(CAST(table2016.percent_fa_region AS NUMERIC),2) AS fa_2016
    FROM table1990
 JOIN table2016
 ON table1990.region = table2016.region;
    WHERE table1990.percent_fa_region > table2016.percent_fa_region;



---3. COUNTRY-LEVEL DETAIL---
---SUCCESS STORIES----


WITH T1 AS
 (SELECT country_name,
 SUM(forest_area_sqkm) forest_area_1
 FROM forestation
 WHERE YEAR = 1990
 GROUP BY country_name,
 forest_area_sqkm),
```

```
 T2 AS
 (SELECT country_name,
 SUM(forest_area_sqkm) forest_area_2
 FROM forestation
 WHERE YEAR = 2016
 GROUP BY country_name,
 forest_area_sqkm)
SELECT f.country_name,
 (f.forest_area_1 - t.forest_area_2) forest_change
FROM T1 f
JOIN T2 t ON f.country_name = t.country_name
ORDER BY forest_change
LIMIT 2;
---------------------
 WITH T1 AS
 (SELECT country_name,
 (SUM(forest_area_sqkm) / SUM(total_area_sq_mi*2.59))*100 percent_forestation_1
 FROM forestation
 WHERE YEAR = 1990
```

```
 GROUP BY country_name,

 forest_area_sqkm),

 T2 AS

 (SELECT country_name,

 (SUM(forest_area_sqkm) / SUM(total_area_sq_mi*2.59))*100 percent_forestation_2

 FROM forestation

 WHERE YEAR = 2016

 GROUP BY country_name,

 forest_area_sqkm)

SELECT f.country_name,

 Round((((f.percent_forestation_1 -

t.percent_forestation_2)/(f.percent_forestation_1))*100)::Numeric, 2) percent_change

FROM T1 f

JOIN T2 t ON f.country_name = t.country_name

ORDER BY percent_change

LIMIT 1;
```

---3.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?---

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

```
WITH table1990 AS (SELECT f.country_code,
                    f.country_name,
                    f.year,
                    f.forest_area_sqkm
                        FROM forest_area f
                WHERE f.year = 1990 AND f.forest_area_sqkm IS NOT NULL AND
f.country_name != 'World'
                ),

    table2016 AS (SELECT f.country_code,
                    f.country_name,
                    f.year,
                    f.forest_area_sqkm
                        FROM forest_area f
                WHERE f.year = 2016 AND f.forest_area_sqkm IS NOT NULL AND
f.country_name != 'World'
                )
```

```
 SELECT table1990.country_code,

        table1990.country_name,

        r.region,

        table1990.forest_area_sqkm AS fa_1990_sqkm,

        table2016.forest_area_sqkm AS fa_2016_sqkm,

        table1990.forest_area_sqkm-table2016.forest_area_sqkm AS diff_fa_sqkm

      FROM table1990

      JOIN table2016

      ON table1990.country_code = table2016.country_code

      AND (table1990.forest_area_sqkm IS NOT NULL AND table2016.forest_area_sqkm IS NOT
NULL)

      JOIN regions r ON table2016.country_code = r.country_code

      ORDER BY 6 DESC

      LIMIT 5;
```

----3.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?---

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

```
WITH table1990 AS (SELECT f.country_code,
                          f.country_name,
                          f.year,
                          f.forest_area_sqkm
                               FROM forest_area f
                     WHERE f.year = 1990 AND f.forest_area_sqkm IS NOT NULL AND
f.country_name != 'World'
                    ),


      table2016 AS (SELECT f.country_code,
                           f.country_name,
                           f.year,
                           f.forest_area_sqkm
                               FROM forest_area f
                     WHERE f.year = 2016 AND f.forest_area_sqkm IS NOT NULL AND
f.country_name != 'World'
                    )


 SELECT table1990.country_code,
        table1990.country_name,
```

```
r.region,

table1990.forest_area_sqkm AS fa_1990_sqkm,

table2016.forest_area_sqkm AS fa_2016_sqkm,

table1990.forest_area_sqkm-table2016.forest_area_sqkm AS diff_fa_sqkm,


ABS(ROUND(CAST(((table2016.forest_area_sqkm-table1990.forest_area_sqkm)/table1990.for
est_area_sqkm*100) AS NUMERIC),2)) AS perc_change

FROM table1990

JOIN table2016

ON table1990.country_code = table2016.country_code

AND (table1990.forest_area_sqkm IS NOT NULL AND table2016.forest_area_sqkm IS NOT
NULL) JOIN regions r ON table2016.country_code = r.country_code

ORDER BY
ROUND(CAST(((table2016.forest_area_sqkm-table1990.forest_area_sqkm)/table1990.forest_a
rea_sqkm*100) AS NUMERIC),2)

LIMIT 5;
```

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

---Table 3.3: Count of Countries Grouped by Forestation Percent Quartiles, 2016:---

```
WITH T1 AS
 (SELECT country_name, year ,
 (SUM(forest_area_sqkm) / SUM(total_area_sq_mi*2.59))*100 percent_forestation
 FROM forestation
 WHERE year = 2016
 GROUP BY country_name, year, forest_area_sqkm)
SELECT Distinct(quartiles), COUNT(country_name)Over(PARTITION BY quartiles)
FROM
 (SELECT country_name,
 CASE
 WHEN percent_forestation<25 THEN '0-25%'
 WHEN percent_forestation>=25
 AND percent_forestation<50 THEN '25%-50%'
 WHEN percent_forestation>=50
 AND percent_forestation<75 THEN '50%-75%'
 ELSE '75%-100%'
 END AS quartiles
 FROM T1
```

```
 WHERE percent_forestation IS NOT NULL

 AND YEAR = 2016) sub;
```

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

---Table 3.4: Top Quartile Countries, 2016:---

```
With table1 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,
                     (f.forest_area_sqkm/(l.total_area_sq_mi*2.59))*100 AS perc_fa
                     FROM forest_area f
                     JOIN land_area l
```

```
                ON f.country_code = l.country_code

                AND (f.country_name != 'World' AND f.forest_area_sqkm IS NOT NULL AND
l.total_area_sq_mi IS NOT NULL)

                AND (f.year=2016 AND l.year = 2016)

                ORDER BY 6 DESC

            ),

    table2 AS (SELECT table1.country_code,

                table1.country_name,

                 table1.year,

                 table1.perc_fa,

                 CASE WHEN table1.perc_fa >= 75 THEN '75%-100%'                    WHEN
table1.perc_fa < 75 AND table1.perc_fa >= 50 THEN '50%-75%'

                    WHEN table1.perc_fa < 50 AND table1.perc_fa >=25 THEN '25%-50%'

                    ELSE '0-25%'

                 END AS percentile

                 FROM table1 ORDER BY 5 DESC

            )

SELECT table2.country_name,

    r.region,

    ROUND(CAST(table2.perc_fa AS NUMERIC),2) AS perc_fa,
```

```
	table2.percentile
	FROM table2
	JOIN regions r
	ON table2.country_code = r.country_code
	WHERE table2.percentile = '75%-100%'
	ORDER BY 1;
```

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

```
With table1 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,
		 (f.forest_area_sqkm/(l.total_area_sq_mi*2.59))*100 AS perc_fa
		 FROM forest_area f
		 JOIN land_area l
		 ON f.country_code = l.country_code
		 AND (f.country_name != 'World' AND f.forest_area_sqkm IS NOT NULL AND
l.total_area_sq_mi IS NOT NULL)
		 AND (f.year=2016 AND l.year = 2016)
```

```
                ORDER BY 6 DESC)

SELECT COUNT(table1.country_name)

    FROM table1

    WHERE table1.perc_fa > (SELECT table1.perc_fa

                            FROM table1

                            WHERE table1.country_name = 'United States');
```