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 41,282,694.90 (sq km) in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 39,958,245.90 (sq km), a loss of 1,324,449.00 (sq km), 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 1,279,999.99 sq km).

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	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 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 (sq km). 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 of America, but it only saw an increase of 79,200 (sq km), much lower than the figure for China.

China and the United States of America 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 5 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	541,510.00 sq km
Indonesia	East Asia & Pacific	282,193.98 sq km
Myanmar	East Asia & Pacific	107,234.00 sq km
Nigeria	Sub-Saharan Africa	106,506.00 sq km
Tanzania	Sub-Saharan Africa	102,320.00 sq km

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 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
1	96
2	72
3	38
4	9

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?
    - *China has done a great job during the time analyzed in order to increase their forest area, it was a huge difference between 2016 and 1990. Not only is it growing economic talking, but also shown to us that is possible to grow thinking about the responsibilities with the environment.*

      On the other hand, Brazil, the biggest country in Latin America & Caribbean region, is demonstrating that they did not have managed well their environmental issues with more than a half-million of forest area (sq km) lost.

      Lastly, Nigeria needs to be monitored more closely, it was the only country that "ranks in the top 5 both in terms of absolute square kilometer decrease in forest as well as the percent decrease in forest area from 1990 to 2016". What we can suggest for them, taking an example of Cabo Verde and Rwanda from the same region, which did a great job increasing their forest percent area by more than 50%, to reverse this situation?
- Which countries should we focus on over others?
    - *Considering the data analyzed, we first need to focus on the top 10 countries appeared above in both lists:*
        Top 5 Amount Decrease in Forest Area by Country, 1990 & 2016
        Top 5 Percent Decrease in Forest Area by Country, 1990 & 2016
      They need to stop losing expressive areas of forest.

6. APPENDIX: SQL Queries Used

```
/*
Student name: Mathaus Vila Nova
LinkedIn: https://www.linkedin.com/in/mathausvilanova/
Course: Udacity SQL Nanodegree Program
*/

/* Introduction */
-- Create a View called "forestation" by joining all three tables -
forest_area, land_area and regions in the workspace.
CREATE VIEW forestation AS (
      SELECT
            fa.country_code,
            fa.country_name,
            fa.year,
            fa.forest_area_sqkm,
            ((la.total_area_sq_mi)*2.59) AS total_area_sqkm,
```

```
            rg1.income_group,
            rg1.region,
            ROUND(CAST(float8
(fa.forest_area_sqkm/((la.total_area_sq_mi*2.59)+0.0010)*100) AS numeric),
4) AS percent_forest_area
      FROM
            forest_area AS fa
            INNER JOIN land_area AS la ON CONCAT(fa.country_code, fa.year)
= CONCAT(la.country_code, la.year)
            INNER JOIN regions AS rg1 ON fa.country_code = rg1.country_code
            --INNER JOIN regions AS rg2 ON la.country_code =
rg2.country_code
);


/* 1. GLOBAL SITUATION */
-- a. What was the total forest area (in sq km) of the world in 1990?
SELECT
      country_code,
      country_name,
      SUM(forest_area_sqkm) AS total_forest_area_sqkm
FROM forestation
WHERE
      country_code='WLD'
      AND year='1990'
GROUP BY country_code, country_name

-- b. What was the total forest area (in sq km) of the world in 2016??
SELECT
      country_code,
      country_name,
      SUM(forest_area_sqkm) AS total_forest_area_sqkm
FROM forestation
WHERE
      country_code='WLD'
      AND year='2016'
GROUP BY country_code, country_name

-- c. What was the change (in sq km) in the forest area of the world from
1990 to 2016?
WITH cte1 AS (
      SELECT
            country_code,
            country_name,
            SUM(forest_area_sqkm) AS total_forest_area_sqkm
      FROM forestation
      WHERE
            country_code='WLD'
            AND year='1990'
```

```
      GROUP BY country_code, country_name
),
cte2 AS (
      SELECT
            country_code,
            country_name,
            SUM(forest_area_sqkm) AS total_forest_area_sqkm
      FROM forestation
      WHERE
            country_code='WLD'
            AND year='2016'
      GROUP BY country_code, country_name
)

SELECT
      ct1.country_code,
      ct1.country_name,
      ct1.total_forest_area_sqkm AS forest_area_sqkm_90,
      ct2.total_forest_area_sqkm AS forest_area_sqkm_16,
      ROUND(CAST(float8 (ct2.total_forest_area_sqkm -
ct1.total_forest_area_sqkm) AS numeric), 2) AS forest_area_change_sqkm, --
between 2016 and 1990
      ROUND(CAST(float8
((ct2.total_forest_area_sqkm/ct1.total_forest_area_sqkm)-1)*(100) AS
numeric), 2) AS forest_area_change_percent -- between 2016 and 1990
FROM
      cte1 AS ct1
      INNER JOIN cte2 AS ct2 ON ct1.country_code = ct2.country_code

-- d. 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 (
      SELECT
            country_code,
            country_name,
            ROUND(CAST(float8 SUM(total_area_sqkm) AS numeric), 2) AS
total_land_area
      FROM forestation
      WHERE year='2016'
      GROUP BY country_code, country_name
) AS tb1
WHERE
      total_land_area < (
            SELECT
                  ((t2.total_forest_area_sqkm -
t1.total_forest_area_sqkm)*(-1))
            FROM (
```

```
                SELECT
                      country_code,
                      country_name,
                      SUM(forest_area_sqkm) AS total_forest_area_sqkm
                FROM forestation
                WHERE
                      country_code='WLD'
                      AND year='1990'
                GROUP BY country_code, country_name
          ) AS t1
                INNER JOIN (
                SELECT
                      country_code,
                      country_name,
                      SUM(forest_area_sqkm) AS total_forest_area_sqkm
                FROM forestation
                WHERE
                      country_code='WLD'
                      AND year='2016'
                GROUP BY country_code, country_name
          ) AS t2 ON t1.country_code = t2.country_code
)
ORDER BY total_land_area DESC
LIMIT 1


/* 2. REGIONAL OUTLOOK */
-- 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?
SELECT
      region,
      ROUND(CAST(float8 (SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100 AS
numeric), 2) AS percent_forest_area
FROM forestation
WHERE
      year='2016'
      --AND region='World'
GROUP BY region
ORDER BY 2
/*
Comments: uncomment the "AND region='World'" in the WHERE clause will give
us the result considering only the World region in 2016.
To check the highest and lowest region, we can run the above query again
and order our result (ORDER BY) by ASC to see the lowest and DESC to see
the highest.
*/

-- 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(CAST(float8 (SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100 AS
numeric), 2) AS percent_forest_area
FROM forestation
WHERE
	year='1990'
	--AND region='World'
GROUP BY region
ORDER BY 2 DESC
/*
Comments: uncomment the "AND region='World'" in the WHERE clause will give
us the result considering only the World region in 1990.
To check the highest and lowest region, we can run the above query again
and order our result (ORDER BY) by ASC to see the lowest and DESC to see
the highest.
*/

-- c. Based on the table you created, which regions of the world DECREASED
in forest area from 1990 to 2016?
WITH cte1 AS (
	SELECT
		region,
		ROUND(CAST(float8
(SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100 AS numeric), 2) AS
percent_forest_area_16
	FROM forestation
	WHERE
		year='2016'
	GROUP BY region
), cte2 AS (
	SELECT
		region,
		ROUND(CAST(float8
(SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100 AS numeric), 2) AS
percent_forest_area_90
	FROM forestation
	WHERE
		year='1990'
	GROUP BY region
)

SELECT
	ct1.region,
	ct2.percent_forest_area_90,
	ct1.percent_forest_area_16,
	CASE
```

```
            WHEN percent_forest_area_16 > percent_forest_area_90 THEN
'INCREASED'
            WHEN percent_forest_area_16 < percent_forest_area_90 THEN
'DECREASED'
            ELSE 'SAME'
      END AS comparison_90_16
FROM
      cte1 AS ct1
      INNER JOIN cte2 AS ct2 ON ct1.region = ct2.region
ORDER BY 4


/* 3. COUNTRY-LEVEL DETAIL */
-- 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?
-- 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 cte1 AS (
      SELECT
            country_code,
            country_name,
            region,
            COALESCE(SUM(forest_area_sqkm), 0.0) AS forest_area_sqkm_90
      FROM forestation
      WHERE
            forest_area_sqkm IS NOT NULL
            AND country_code != 'WLD'
            AND year='1990'
      GROUP BY country_code, country_name, region
), cte2 AS (
      SELECT
            country_code,
            country_name,
            region,
            COALESCE(SUM(forest_area_sqkm), 0.0) AS forest_area_sqkm_16
      FROM forestation
      WHERE
            forest_area_sqkm IS NOT NULL
            AND country_code != 'WLD'
            AND year='2016'
      GROUP BY country_code, country_name, region
)

SELECT
      ct1.country_code,
      ct1.country_name,
      ct1.region,
      ct1.forest_area_sqkm_90,
```

```
      ct2.forest_area_sqkm_16,
      CASE
            WHEN ct1.forest_area_sqkm_90 > 0 THEN ROUND(CAST(float8
(ct2.forest_area_sqkm_16-ct1.forest_area_sqkm_90) AS numeric), 2)
            ELSE 0.0
      END AS forest_diff_sqkm,
      CASE
            WHEN ct1.forest_area_sqkm_90 > 0 THEN ROUND(CAST(float8
((ct2.forest_area_sqkm_16/(ct1.forest_area_sqkm_90+0.001))-1)*100 AS
numeric), 2)
            ELSE 0.0
      END AS forest_diff_sqkm_percent
FROM
      cte1 AS ct1
      INNER JOIN cte2 AS ct2 ON CONCAT(ct1.country_code, ct1.country_name)
= CONCAT(ct2.country_code, ct2.country_name)
ORDER BY 7
LIMIT 5

/*
Comments: to change between a. and b. question's answer, you must need to
change the ORDER BY clause.
ORDER BY 6 ASC if you want to see the largest amount decrease and ORDER BY
7 ASC if you want to see the largest percent decrease.
*/

-- c. If countries were grouped by percent forestation in quartiles, which
group had the most countries in it in 2016?
SELECT
      quartile,
      COUNT(1) count_of_countries
FROM (
      SELECT
            CASE
                  WHEN percent_forest_area BETWEEN 0 AND 25 THEN '1'
                  WHEN percent_forest_area BETWEEN 25 AND 50 THEN '2'
                  WHEN percent_forest_area BETWEEN 50 AND 75 THEN '3'
                  ELSE '4'
            END AS quartile
      FROM forestation
      WHERE
            country_code!='WLD'
            AND total_area_sqkm > 0
            AND year='2016'
) AS t1
GROUP BY quartile
ORDER BY 2 DESC

--d. List all of the countries that were in the 4th quartile (percent
```

```
forest > 75%) in 2016.
SELECT
	country_name,
	region,
	ROUND(CAST(float8 (percent_forest_area) AS numeric), 2) AS
percent_forest_area
FROM forestation
WHERE
	country_code!='WLD'
	AND total_area_sqkm > 0
	AND year='2016'
	AND percent_forest_area > 75
ORDER BY 3 DESC

--e. How many countries had a percent forestation higher than the United
States in 2016?
WITH cte AS (
	SELECT
		country_code,
		country_name,
		region,
		year,
		SUM(total_area_sqkm) AS land_area_sqkm,
		SUM(forest_area_sqkm) AS forest_area_sqkm,
		CASE
			WHEN ROUND(CAST(float8
(SUM(forest_area_sqkm)/(SUM(total_area_sqkm)+0.0001))*100 AS numeric), 2) >
100 THEN CAST(100 AS float)
			ELSE ROUND(CAST(float8
(SUM(forest_area_sqkm)/(SUM(total_area_sqkm)+0.0001))*100 AS numeric), 2)
		END AS forest_percent_area
	FROM forestation
	WHERE
		forest_area_sqkm IS NOT NULL
		AND country_code != 'WLD'
		AND year='2016'
	GROUP BY country_code, country_name, year, region
)

SELECT
	COUNT(1) AS count_of_countries
FROM cte
WHERE
	forest_percent_area > (
			SELECT
				forest_percent_area
			FROM cte
			WHERE
```

```
                country_code='USA'
                AND year='2016')
```