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 square kilometers (sq/km) in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 39,958,245.9 sq/km, a loss of 1,324,449.0 sq/km, a decrease of 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 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 Middle East & North Africa, with 2.07% forestation.

In 1990, the percentage 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 50.03%, 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
World	32.42	31.38
Sub-Saharan Africa	30.67	28.79
South Asia	16.51	17.51
North America	35.65	36.04
Middle East & North Africa	1.78	2.07
Latin America & Caribbean	51.03	46.16
Europe & Central Asia	37.28	38.04
East Asia & Pacific	25.78	26.36

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 (dropped from 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 33.55%. 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 2.62%, much lower than the figure for China.

China and the 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's forest area increased 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	SQ/KM Forest Area Change
Brazil	Latin America & Caribbean	541,510
Indonesia	East Asia & Pacific	282,194
Myanmar	East Asia & Pacific	107,234
Nigeria	Sub-Saharan Africa	106,506
Tanzania	Sub-Saharan Africa	102,320

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
Q1	85
Q2	72
Q3	38
Q4	9

The largest number of countries in 2016 were found in the first quartile. The count for the number of countries in table 3.3 is less than the 217 countries in the data for 2016, due to 13 countries not having any value recorded for their forest area in that year.

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
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

One aspect of the data that hasn't been considered in this report is the association of income status and deforestation by country. The following list displays the four income groups found in the deforestation dataset:

Table 5.1: Counts of countries by Income Group

Income Group	Count	PCT of Count
Upper middle income	56	25.80
Lower middle income	47	21.66
High income	80	36.86
Low income	34	15.68

The majority of the countries fall into either the High income or Upper middle income categories.

Tables 5.2 and 5.3 below, reproduce tables 3.1 and 3.2 above, with the addition of the income level defined in the data. The first table focuses on the absolute square kilometer change in forest area from 1990 to 2016. Listed in descending order by the forest area change.
Table 5.2: Top 5 Amount Decrease in Forest Area by Country, 1990 & 2016:

Country	Income Group	SQ/KM Forest Area Change
Brazil	Upper middle income	541,510
Indonesia	Lower middle income	282,194
Myanmar	Lower middle income	107,234
Nigeria	Lower middle income	106,506
Tanzania	Lower middle income	102,320

The second table lists countries by the percent forest area change, in descending order, from 1990 to 2016.

Table 5.3: Top 5 Percent Decrease in Forest Area by Country, 1990 & 2016:

Country	Income Group	Pct Forest Area Change
Togo	Low income	75.45
Nigeria	Lower middle income	61.80
Uganda	Low income	59.13
Mauritania	Lower middle income	46.75
Honduras	Lower middle income	45.03

It is interesting to note that, only one out of the ten countries listed in these two tables, is considered to be in the higher income category, there are zero countries in the High income group, and the rest are from the lower income groups.

The following tables, 5.4 and 5.5, look at the distribution of change in forest area by square kilometer and by percent, from 1990 to 2016, when the countries are grouped by their income group.

It is interesting to see that the highest income group gained forest area, while the other categories had a decrease in forest area. This may be related to the stage of a country in its

development lifecycle, but that information is not part of the data set. It is possibly a topic for additional study.

Table 5.4: Change in Forest Area Square Kilometers by Income Group, 1990 & 2016:

Income Group	SQ/KM Forest 1990	SQ/KM Forest 2016	SQ/KM Forest Area Change
Upper middle income	20,053,166	19,878,288	-174,877
Lower middle income	6,273,591	5,820,013	-453,578
High income	10,316,559	10,435,350	118,791
Low income	4,090,460	3,733,535	-356,924

Table 5.5: Change in Forest Area Percent by Income Group, 1990 & 2016:

Income Group	Pct Forest 1990	Pct Forest 2016	Pct Forest Area Change
Upper middle income	36.08	35.79	-0.29
Lower middle income	31.80	29.54	-2.26
High income	27.33	27.57	0.24
Low income	28.74	26.23	-2.51

One limitation of the data set is that a country's Income group is static, therefore we are not able to measure the change in a country's income group over a given period and see how well it correlates with a country's change in forestation.

It would facilitate additional analysis, if the dataset were to include information regarding a country's Income group for each year, in addition to an indicator of it stage in industrial development lifecycle.

APPENDIX: SQL queries used

NOTE:
For easier viewing or testing, these queries can be found at:
https://github.com/brian-sigurdson/udacity-courses/tree/main/nano_sql/0_intro_to_sql/project.

See:

- create_forestation_view.sql
- global_situation.sql
- regional_outlook.sql
- country_level_detail.sql
- recommendations.sql

Create Project View

```
-- view for project

CREATE OR replace VIEW forestation AS
SELECT r.country_name,
    r.country_code,
    r.region,
    r.income_group,
    f.year,
    f.forest_area_sqkm,
    f.forest_area_sqkm / 2.59 AS forest_area_sq_mi,
    la.total_area_sq_mi,
    la.total_area_sq_mi * 2.59 AS total_area_sqkm,
    ROUND( (f.forest_area_sqkm / (la.total_area_sq_mi * 2.59) * 100)::NUMERIC, 2) AS
percent_forest
FROM regions r
INNER JOIN forest_area f ON r.country_code = f.country_code
INNER JOIN land_area la ON f.country_code = la.country_code AND f.year = la.year
ORDER BY year, country_name
```

Global Situation

```
-- These queries are associated WITH the questions presented in the Global Situation section.

-- 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 forest_area_sqkm AS forest_area_sqkm_1990
FROM forestation
WHERE year = 1990 AND country_code = 'WLD';

-- 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 forest_area_sqkm AS forest_area_sqkm_2016
FROM forestation
WHERE year = 2016 AND country_code = 'WLD';

-- c. What was the change (in sq km) in the forest area of the world FROM 1990 to 2016?
WITH sqkm_1990 AS
   (
      SELECT country_code, forest_area_sqkm AS forest_area_sqkm_1990
      FROM forestation
      WHERE year = 1990 AND country_code = 'WLD'
   ),
   sqkm_2016 AS
   (
      SELECT country_code, forest_area_sqkm AS forest_area_sqkm_2016
      FROM forestation
      WHERE year = 2016 AND country_code = 'WLD'
   )
SELECT sqkm_1990.country_code,
      forest_area_sqkm_1990,
      forest_area_sqkm_2016,
      forest_area_sqkm_2016 - forest_area_sqkm_1990 AS sqkm_change
FROM sqkm_1990
INNER JOIN sqkm_2016 ON sqkm_1990.country_code = sqkm_2016.country_code;

-- d. What was the percent change in forest area of the world between 1990 AND 2016?
WITH sqkm_1990 AS
   (
      SELECT country_code, forest_area_sqkm AS forest_area_sqkm_1990
      FROM forestation
      WHERE year = 1990 AND country_code = 'WLD'
   ),
   sqkm_2016 AS
   (
      SELECT country_code, forest_area_sqkm AS forest_area_sqkm_2016
      FROM forestation
```

```
        WHERE year = 2016 AND country_code = 'WLD'
    )
SELECT sqkm_1990.country_code,
       forest_area_sqkm_1990,
       forest_area_sqkm_2016,
       forest_area_sqkm_2016 - forest_area_sqkm_1990 AS sqkm_change,
       ROUND((((forest_area_sqkm_2016 - forest_area_sqkm_1990) /
forest_area_sqkm_1990)::NUMERIC * 100),2) AS percentage_change
FROM sqkm_1990
INNER JOIN sqkm_2016 ON sqkm_1990.country_code = sqkm_2016.country_code;

-- 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?
-- Find the first country whose land area is more than the forest area lost between 1990 AND
2016, when the land areas
-- BY country are in descending ORDER.
SELECT country_name, ROUND(total_area_sqkm::NUMERIC,2)
FROM forestation
WHERE year = 2016 AND country_code != 'WLD' AND total_area_sqkm <
                (
                   -- calculate the forest area lost between 1990 AND 2016
                   SELECT abs(sqkm_change)
                   FROM (
                         WITH sqkm_1990 AS
                               (
                                  SELECT country_code, forest_area_sqkm AS
forest_area_sqkm_1990
                                  FROM forestation
                                  WHERE year = 1990
                                    AND country_code = 'WLD'
                               ),
                            sqkm_2016 AS
                               (
                                  SELECT country_code, forest_area_sqkm AS
forest_area_sqkm_2016
                                  FROM forestation
                                  WHERE year = 2016
                                    AND country_code = 'WLD'
                               )
                         SELECT forest_area_sqkm_2016 - forest_area_sqkm_1990 AS
sqkm_change
                         FROM sqkm_1990
```

```
                                INNER JOIN sqkm_2016 ON sqkm_1990.country_code =
sqkm_2016.country_code
                            ) AS sub
                    )
ORDER BY total_area_sqkm desc
LIMIT 1
```

Regional Outlook

```
-- This query is associated with the questions presented in the REGIONAL OUTLOOK section.

-- 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. (Note that 1 sq mi =
2.59 sq km).
--      Based ON the table you created, answer the following questions.
WITH sqkm_1990 AS
   (
      SELECT region,
            SUM(forest_area_sqkm) AS sum_forest_area_sqkm_1990,
            SUM(total_area_sqkm) AS sum_land_area_sqkm_1990,
            ROUND((SUM(forest_area_sqkm) / SUM(total_area_sqkm))::NUMERIC * 100,2) AS
percentage_forest_area_1990
      FROM forestation
      WHERE year = 1990
      GROUP BY 1
   ),
   sqkm_2016 AS
   (
      SELECT region,
            SUM(forest_area_sqkm) AS sum_forest_area_sqkm_2016,
            SUM(total_area_sqkm) AS sum_land_area_sqkm_2016,
            ROUND((SUM(forest_area_sqkm) / SUM(total_area_sqkm))::NUMERIC * 100,2) AS
percentage_forest_area_2016
      FROM forestation
      WHERE year = 2016
      GROUP BY 1
   )
SELECT sqkm_1990.region,
      percentage_forest_area_1990,
      percentage_forest_area_2016,
      percentage_forest_area_2016 - percentage_forest_area_1990 AS
percentage_forest_area_change
```

```
FROM sqkm_1990
INNER JOIN sqkm_2016 ON sqkm_1990.region = sqkm_2016.region
-- ORDER BY sqkm_forest_area_change
ORDER BY percentage_forest_area_1990


-- a. What was the percent forest of the entire world in 2016?
-- 31.38
-- Which region had the HIGHEST percent forest in 2016, AND which had the LOWEST, to 2 decimal places?
-- Highest = Latin America & Caribbean, Lowest = Middle East & North Africa

-- b. What was the percent forest of the entire world in 1990?
-- 32.42
-- Which region had the HIGHEST percent forest in 1990, AND which had the LOWEST, to 2 decimal places?
-- Highest = Latin America & Caribbean, Lowest = Middle East & North Africa

-- c. Based ON the table you created, which regions of the world DECREASED in forest area FROM 1990 to 2016?
-- Latin America & Caribbean, Sub-Saharan Africa
```

Country Level Detail

```
-- These queries are associated WITH the questions presented in the COUNTRY-LEVEL DETAIL section.

-- The first paragraph of part 3 of the document, COUNTRY-LEVEL DETAIL, states information about success stories,
-- but were not included in the preparation questions below, so I'm creating it now.

-- success stories
-- Which country increased forest area FROM 1990 to 2016?
-- What was the percentage change?
-- Which is the country WITH the next largest increase in forest area FROM 1990 to 2016?
-- What was the percentage change?
WITH sqkm_1990 AS
    (
        SELECT country_name, forest_area_sqkm AS forest_area_sqkm_1990
        FROM forestation
        WHERE year = 1990 AND country_name != 'World'
    ),
```

```
   sqkm_2016 AS
   (
      SELECT country_name, forest_area_sqkm AS forest_area_sqkm_2016
      FROM forestation
      WHERE year = 2016 AND country_name != 'World'
   )
SELECT sqkm_1990.country_name,
      ROUND((forest_area_sqkm_2016 - forest_area_sqkm_1990)::NUMERIC,2) AS
sqkm_change,
      ROUND(((forest_area_sqkm_2016 -
forest_area_sqkm_1990)/(forest_area_sqkm_1990))::NUMERIC * 100,2) AS percent_change
FROM sqkm_1990
INNER JOIN sqkm_2016 ON sqkm_1990.country_name = sqkm_2016.country_name
WHERE (forest_area_sqkm_2016 - forest_area_sqkm_1990) is not null
ORDER BY percent_change desc
LIMIT 2;

-- 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?
WITH sqkm_1990 AS
   (
      SELECT country_name, forest_area_sqkm AS forest_area_sqkm_1990
      FROM forestation
      WHERE year = 1990 AND country_name != 'World'
   ),
   sqkm_2016 AS
   (
      SELECT country_name, forest_area_sqkm AS forest_area_sqkm_2016
      FROM forestation
      WHERE year = 2016 AND country_name != 'World'
   )
SELECT sqkm_1990.country_name, ROUND((forest_area_sqkm_2016 -
forest_area_sqkm_1990)::NUMERIC,2) AS sqkm_change
FROM sqkm_1990
INNER JOIN sqkm_2016 ON sqkm_1990.country_name = sqkm_2016.country_name
ORDER BY sqkm_change
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 sqkm_1990 AS
   (
      SELECT country_name, region, forest_area_sqkm AS forest_area_sqkm_1990
```

```
    FROM forestation
    WHERE year = 1990 AND country_name != 'World'
  ),
  sqkm_2016 AS
  (
    SELECT country_name, region, forest_area_sqkm AS forest_area_sqkm_2016
    FROM forestation
    WHERE year = 2016 AND country_name != 'World'
  )
SELECT sqkm_1990.country_name, sqkm_1990.region,
    ROUND((forest_area_sqkm_2016 - forest_area_sqkm_1990)::NUMERIC,2) AS
sqkm_change,
    ROUND(((forest_area_sqkm_2016 -
forest_area_sqkm_1990)/(forest_area_sqkm_1990))::NUMERIC * 100,2) AS percent_change
FROM sqkm_1990
INNER JOIN sqkm_2016 ON sqkm_1990.country_name = sqkm_2016.country_name
WHERE (forest_area_sqkm_2016 - forest_area_sqkm_1990) is not null
-- ORDER BY sqkm_change
ORDER BY percent_change
LIMIT 5;

-- c. If countries were grouped BY percent forestation in quartiles, which GROUP had the most
countries in it in 2016?
-- NOTE:  Per the grading rubric, we need to implement a case statement for this solution.
WITH quartile_sums_2016 AS
  (
    SELECT
      case
        when percent_forest <= 25.00 then 'Q1'
        when percent_forest > 25.00 AND percent_forest <= 50.00 then 'Q2'
        when percent_forest > 50.00 AND percent_forest <= 75.00 then 'Q3'
        when percent_forest > 75.00 then 'Q4'
      END AS quartile
    FROM forestation
    WHERE year = 2016 AND country_code != 'WLD' AND percent_forest IS NOT NULL
  )
SELECT quartile, count(*) as quartile_count
FROM quartile_sums_2016
group by 1
order by 1;

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

```
FROM forestation
WHERE year = 2016 AND percent_forest > 75.00;

-- e. How many countries had a percent forestation higher than the United States in 2016?
-- NOTE:  This is a self-INNER JOIN query.
SELECT COUNT(*)
FROM forestation a
INNER JOIN forestation b ON a.year = b.year AND a.country_code = b.country_code
WHERE a.year = 2016 AND a.percent_forest >
                    (
                       SELECT percent_forest
                       FROM forestation
                       WHERE year = 2016 AND country_code = 'USA'
                    )
```

Recommendations

```
-- These queries are associated with the RECOMMENDATIONS section.

-- Income Group: count, forest area, land area, percent forest
WITH tbl_1990 AS
   (
      SELECT income_group,
         COUNT(*) AS COUNT,
         ROUND(SUM(forest_area_sqkm)::NUMERIC,2) AS ttl_forest_area,
         ROUND(SUM(total_area_sqkm)::NUMERIC,2) AS ttl_land_area,
         ROUND((SUM(forest_area_sqkm) / SUM(total_area_sqkm))::NUMERIC * 100, 2) AS
pct_forest
      FROM forestation
      WHERE year = 1990 AND country_code != 'WLD'
      GROUP BY 1
   ),
 tbl_2016 AS
    (
       SELECT income_group,
            COUNT(*) AS COUNT,
            ROUND(SUM(forest_area_sqkm)::NUMERIC,2) AS ttl_forest_area,
            ROUND(SUM(total_area_sqkm)::NUMERIC,2) AS ttl_land_area,
            ROUND((SUM(forest_area_sqkm) / SUM(total_area_sqkm))::NUMERIC * 100, 2) AS
pct_forest
       FROM forestation
       WHERE year = 2016 AND country_code != 'WLD'
       GROUP BY 1
```

```
    )
SELECT tbl_1990.income_group, tbl_2016.COUNT,
    tbl_1990.ttl_forest_area AS ttl_forest_area_1990,
    tbl_2016.ttl_forest_area AS ttl_forest_area_2016,
    tbl_2016.ttl_forest_area - tbl_1990.ttl_forest_area AS ttl_forest_area_change,
    tbl_1990.pct_forest AS pct_forest_1990,
    tbl_2016.pct_forest AS pct_forest_2016,
    tbl_2016.pct_forest - tbl_1990.pct_forest AS pct_change
FROM tbl_1990
INNER JOIN tbl_2016 ON tbl_1990.income_group = tbl_2016.income_group
```