# Contents

# Report for Forest Query into Global Deforestation, 1990 to 2016

Forest Query 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 Forest Query 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 sq km _ in 1990. As of 2016, the most recent year for which data was available, that number had fallen to __ 39958245.9 sq km _, a loss of __ 1324449 sq km__, or __ 3.20824258980244__%.

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 __ 1279999.9891 sq km__).

## 2. REGIONAL OUTLOOK

In 2016, the percent of the total land area of the world designated as forest was __ 31.3755709643095_%_. The region with the highest relative forestation was__ Latin America & Caribbean__, with __ 46.1620721996047__%, and the region with the lowest relative forestation was __ Middle East & North Africa__, with __ 2.06826486871501__% forestation.

In 1990, the percent of the total land area of the world designated as forest was __ 32.4222035575689__. The region with the highest relative forestation was Latin America & Caribbean__, with __ 51.0299798667514__%, and the region with the lowest relative forestation was __ Middle East & North Africa__, with __ 1.77524062469353__% forestation.

Table 2.1: Percent Forest Area by Region, 1990 & 2016:

| Region | 1990 Forest Percentage | 2016 Forest Percentage |
|---|---|---|
| Latin America & Caribbean | 51.0299798667514 % | 46.1620721996047 % |
| Sub-Saharan Africa | 30.6741454610006 % | 28.7881883550464 % |

The only regions of the world that decreased in percent forest area from 1990 to 2016 were __ Latin America & Caribbean __ (dropped from __ 51.0299798667514 __% to __ 46.1620721996047 __%) and __ Sub-Saharan Africa __ (__ 30.6741454610006 __% to __ 28.7881883550464 __%). 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 __ 32.4222035575689 __% to __ 31.3755709643095 __%.

# 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 __ 527229.062 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 _, but it only saw an increase of _ 2.6186146470491 % _, much lower than the figure for __ 79200 __.

__ Russian Federation __ and __ China __ 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 213.664588870028 % 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 sq km |
| Indonesia | East Asia & Pacific | 282193.9844 sq km |
| Myanmar | East Asia & Pacific | 107234.0039 sq km |
| Nigeria | Sub-Saharan Africa | 106506.00098 sq km |
| Tanzania | Sub-Saharan Africa | 102320 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 | 75.4452559270073 % |
| Nigeria | Sub-Saharan Africa | 61.7999309388418 % |
| Uganda | Sub-Saharan Africa | 59.1286034729531 % |
| Mauritania | Sub-Saharan Africa | 46.7469879518072 % |
| Honduras | Latin America & Caribbean | 45.0344149459194 % |

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 |
|---|---|
| 1st Quartile | 85 |
| 2nd Quartile | 73 |
| 3rd Quartile | 38 |
| 4th Quartile | 9 |

The largest number of countries in 2016 were found in the _____1_____ 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 % |

# 4. RECOMMENDATIONS

*Write out a set of recommendations as an analyst on the Forest Query team.*

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

*I've learned that forested land areas are increasing each year within most of the countries.*

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

*China is a huge country but, their forested land area has been increasing significantly more than any other countries. Forest land area in China was over 20% in recent years. Since the late 1990s, China has planted more than 40 million hectares of forest every year while naturally regenerated forests increased by 22 million hectares.*

# 5. Appendix: SQL queries used

## Global Situation

# 1. Create a **View** called **"forestation"** by joining all three tables - **forest_area, land_area** and **regions** in the workspace.
# 2. The **forest_area** and **land_area** tables *join* on both **country_code** AND **year**. #
3. The **regions** table joins these based on only **country_code**
# 4. In the 'forestation' View, include the following:
# **All of the columns of the origin tables**
# A **new column** that provides the **percent of the land area that is designated as forest**.

```sql
CREATE OR REPLACE 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,
       ((f.forest_area_sqkm / (l.total_area_sq_mi*2.59)) * 100) AS
percent_of_the_land_area_as_forest,
       r.region,
       r.income_group
FROM forest_area f
JOIN land_area l
ON f.country_code=l.country_code AND f.year=l.year
JOIN regions r
ON r.country_code=l.country_code;
```

# This produces the total forest area of the World in 1990.

```sql
SELECT region,
       year,
       forest_area_sqkm
FROM forestation
WHERE region='World' AND year=1990;
```

# This produces the total forest area of the World in 2016.

```sql
SELECT region,
       year,
       forest_area_sqkm
FROM forestation
WHERE region='World' AND year=2016;
```

# What was the change (in sq km) in the forest area of the world from 1990 to 2016?
# The value 1324449 produced as forest sq km decrease.

```sql
SELECT year,
       forest_area_sqkm,
       LEAD(forest_area_sqkm) OVER (ORDER BY year)-forest_area_sqkm AS
change
FROM forestation
WHERE region='World' AND (year=1990 OR year=2016);
```

# What was the percent change in forest area of the world between 1990 and 2016? # This produces the forest area values in 2016 and 1990, the forest loss, and the percent change between 1990 and 2016.

```sql
WITH forest_1990 AS ( SELECT year,
                             region,
                             forest_area_sqkm,
                             percent_of_the_land_area_as_forest AS percent
                FROM forestation
                WHERE region='World' AND year=1990
),
forest_2016 AS ( SELECT year,
                        region,
                        forest_area_sqkm,
                        percent_of_the_land_area_as_forest AS percent
                FROM forestation
                WHERE region='World' AND year=2016
)
```

```sql
SELECT *, (forest_1990.forest_area_sqkm-forest_2016.forest_area_sqkm) As
loss,
((forest_1990.forest_area_sqkm -
forest_2016.forest_area_sqkm)/forest_1990.forest_area_sqkm)*100 As
percent_decrease FROM forest_1990
JOIN forest_2016
ON forest_1990.region=forest_2016.region
```

# This produces Peru as the country with the total land area closet to the amount of forest area lost between 1990 and 2016.

```sql
WITH forest_1990 AS ( SELECT year,
                             country_name,
                             region,
                             forest_area_sqkm,
                             total_area_sqkm,
                             percent_of_the_land_area_as_forest AS percent
                      FROM forestation
                      WHERE year=1990
),
forest_2016 AS ( SELECT year,
                        country_name,
                        region,
                        forest_area_sqkm,
                        total_area_sqkm,
                        percent_of_the_land_area_as_forest AS percent
                 FROM forestation
                 WHERE year=2016
)
```

```sql
SELECT forest_2016.country_name, forest_2016.total_area_sqkm
FROM forest_2016
JOIN forest_1990
ON forest_2016.region=forest_1990.region
WHERE forest_2016.total_area_sqkm <=
(SELECT (forest_1990.forest_area_sqkm-forest_2016.forest_area_sqkm)
As loss_forest
FROM forest_1990
JOIN forest_2016
ON forest_1990.region=forest_2016.region
WHERE forest_2016.region='World')
AND forest_2016.region!='World'
GROUP BY 1,2
ORDER BY 2 DESC
```

# Regional Outlook

# 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)

```sql
WITH forest AS
(
    SELECT year,
           region,
           (SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100 AS
percent_forest_area
    FROM forestation
    WHERE (year=1990 OR year=2016)
    GROUP BY 1,2
)
```

```sql
SELECT *
FROM forest
```

# This returns the percent of the forest area of the entire World in 2016 as 31.38

```sql
SELECT percent_forest_area
FROM forest
WHERE year=2016 AND region='World'
```

# This produces Latin America & Caribbean as the region with the highest forest percentage in 2016.

```sql
SELECT region,percent_forest_area
FROM forest
WHERE year=2016 AND region != 'World'
ORDER BY percent_forest_area DESC
LIMIT 1
```

# This produces Middle East & North Africa as the region with the lowest forest percentage in 2016.

```
SELECT region, percent_forest_area
FROM forest
WHERE year=2016 AND region != 'World'
ORDER BY percent_forest_area
LIMIT 1
```

# This returns the percent of the forest area of the entire World in 1990 as 32.42

```
SELECT percent_forest_area
FROM forest
WHERE year=1990 AND region='World'
```

# This produces Latin America & Caribbean as the region with the highest forest percentage in 1990.

```
SELECT region, percent_forest_area
FROM forest
WHERE year=1990 AND region != 'World'
ORDER BY percent_forest_area DESC
LIMIT 1
```

# This produces Middle East & North Africa as the region with the lowest forest percentage in 1990.

```
SELECT region, percent_forest_area
FROM forest
WHERE year=1990 AND region != 'World'
ORDER BY percent_forest_area
LIMIT 1
```

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

# Only Sub-Saharan Africa and Latin America & Caribbean regions decreased in forest area between 1990 and 2016.

```sql
WITH forest_1990 AS (
    SELECT year, region,
            (SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100 AS
percent_per_region
    FROM forestation
    WHERE region!='World' AND year=1990
    GROUP BY 1,2
),
forest_2016 AS (
    SELECT year, region,
            (SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100 AS
percent_per_region
    FROM forestation
    WHERE region!='World' AND year=2016
    GROUP BY 1,2
)
```

```sql
SELECT region,
        percent_1990,
        percent_2016,
        (percent_1990-percent_2016) AS diff
FROM (SELECT forest_2016.region AS region,
            forest_1990.percent_per_region AS "percent_1990",
            forest_2016.percent_per_region AS "percent_2016"
FROM forest_1990
JOIN forest_2016
ON forest_1990.region=forest_2016.region) sub
ORDER BY diff DESC;
```

# the percent forest area of the world decreased over this time period from __32.4222035575689__% to __31.3755709643095__%.

```sql
SELECT year,
       region,
       (SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100 AS
percent_per_region FROM forestation
WHERE region='World' AND (year=1990 OR year=2016)
GROUP BY 1,2
```

# Country-Level Detail

## SUCCESS STORIES

# This produces China and United States as the top two countries in terms of forest area increase in terms of sq km.

```sql
WITH forest_1990 AS(
    SELECT country_name,
           forest_area_sqkm,
           SUM(forest_area_sqkm) forest_area,
           percent_of_the_land_area_as_forest
    FROM forestation
    WHERE forest_area_sqkm IS NOT NULL
    AND country_name != 'World'
    AND year=1990 GROUP BY 1,2,4
    ORDER BY 2 DESC),

forest_2016 AS(
    SELECT country_name,
           forest_area_sqkm,
           SUM(forest_area_sqkm) forest_area,
           percent_of_the_land_area_as_forest
    FROM forestation
    WHERE forest_area_sqkm IS NOT NULL
    AND country_name != 'World' AND year=2016 GROUP BY 1,2,4
    ORDER BY 2 DESC)
```

```sql
SELECT forest_1990.country_name,
       forest_1990.forest_area AS "forest_1990",
       forest_2016.forest_area AS "forest_2016",
       (forest_1990.forest_area_sqkm-forest_2016.forest_area_sqkm) AS
difference,

((forest_1990.forest_area_sqkmforest_2016.forest_area_sqkm)/forest_1990.for
est_area_sqkm)*100 As percent_decreased
FROM forest_1990
JOIN forest_2016
ON forest_1990.country_name = forest_2016.country_name
ORDER BY difference;
```

# large countries in total land area

```sql
SELECT *
FROM
(SELECT country_name,
        SUM(total_area_sqkm) AS total_area
 FROM forestation
 GROUP BY 1
 ORDER BY 2 DESC) sub
WHERE total_area IS NOT NULL AND country_name != 'World'
```

# 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 213.664588870028 % from 1990 to 2016.

# This produces Iceland as the top country in terms of percentage change increase in forest area.

```sql
WITH forest_1990 AS(
    SELECT country_name,
            region,
            forest_area_sqkm
    FROM forestation
    WHERE forest_area_sqkm IS NOT NULL
    AND country_name != 'World'
    AND year=1990
    GROUP BY 1,2,3
    ORDER BY 3 DESC),

forest_2016 AS(
    SELECT country_name,
            region,
            forest_area_sqkm
    FROM forestation
    WHERE forest_area_sqkm IS NOT NULL
    AND country_name != 'World'
    AND year=2016
    GROUP BY 1,2,3
    ORDER BY 3 DESC)
```

```sql
SELECT forest_1990.country_name,
        forest_1990.forest_area_sqkm AS forest_1990,
        forest_2016.forest_area_sqkm AS forest_2016,
((forest_1990.forest_area_sqkmforest_2016.forest_area_sqkm)/forest_1990.for
est_area_sqkm)*100 AS percent_decrease
FROM forest_1990
JOIN forest_2016
ON forest_1990.country_name=forest_2016.country_name
ORDER BY percent_decrease
```

## LARGEST CONCERNS

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

# The top five countries in terms of sq km forestation decrease are Brazil, Indonesia, Myanmar, Nigeria, and Tanzania.

```sql
WITH forest_1990 AS(
    SELECT country_name,
           forest_area_sqkm,
           region
    FROM forestation
    WHERE year=1990
    ORDER BY country_name
),
forest_2016 AS(
    SELECT country_name,
           forest_area_sqkm,
           region
    FROM forestation
    WHERE year=2016
    ORDER BY country_name
)
```

```sql
SELECT *
FROM
 (SELECT forest_1990.country_name,
         forest_1990.region,
         forest_1990.forest_area_sqkm as before,
         forest_2016.forest_area_sqkm as later,
         forest_1990.forest_area_sqkm-forest_2016.forest_area_sqkm as
difference
  FROM forest_1990
  JOIN forest_2016
  ON forest_1990.country_name=forest_2016.country_name
  WHERE forest_1990.country_name != 'World') sub
WHERE difference IS NOT NULL
ORDER BY difference DESC
LIMIT 5;
```

\#  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?

\# The top five countries in terms of percent decrease in forestation are Togo, Nigeria,
Uganda, Mauritania, and Honduras.

Q. How to calculate the percent decrease then you can use the formula? (Change
in forest_area / Actual forest_area for 1990 )* 100

```sql
WITH forest_1990 AS(
    SELECT country_name,
           forest_area_sqkm,
           region,
           percent_of_the_land_area_as_forest
    FROM forestation
    WHERE year=1990
    ORDER BY country_name
),
forest_2016 AS(
    SELECT country_name,
           forest_area_sqkm,
           region,
           percent_of_the_land_area_as_forest
    FROM forestation
    WHERE year=2016
    ORDER BY country_name
)
```

```sql
SELECT *
FROM
   (SELECT forest_1990.country_name,
           forest_1990.region,
           forest_1990.percent_of_the_land_area_as_forest as before,
           forest_2016.percent_of_the_land_area_as_forest as later,
((forest_1990.forest_area_sqkmforest_2016.forest_area_sqkm)/forest_19
90.forest_area_sqkm)*100 as percent_decrease
   FROM forest_1990
   JOIN forest_2016
   ON forest_1990.country_name=forest_2016.country_name) sub
WHERE percent_decrease IS NOT NULL
ORDER BY percent_decrease DESC
LIMIT 5;
```

QUARTILES

The countries should be grouped according to their percentage forest area. For example, countries having >75% forest area in the 4th quartile, those having >50% but <=75% are in the 3rd quartile and so on.
The NTILE function groups them on percentiles, which is not desired

# If countries were grouped by percent forestation in quartiles, which group had the most countries in it in 2016?
Check each group with this WHERE clause
WHERE qual.forest_quartile = '?'

```
WITH qual AS (
    SELECT country_name,
           region,
           ROUND(percent_of_the_land_area_as_forest::DECIMAL,2) AS
percent_forest,
           CASE WHEN percent_of_the_land_area_as_forest <=25 THEN '1'
                WHEN percent_of_the_land_area_as_forest <=50 THEN '2'
                WHEN percent_of_the_land_area_as_forest <=75 THEN '3'
                ELSE '4' END AS forest_quartile
    FROM forestation
    WHERE year = 2016
    AND forest_area_sqkm IS NOT NULL
    AND total_area_sqkm IS NOT NULL
    ORDER BY percent_of_the_land_area_as_forest)
```

```
SELECT qual.country_name,
       qual.region,
       qual.percent_forest,qual.forest_quartile
FROM qual
ORDER BY 3 DESC
```

# To get the actual count for the respective columns,

```
SELECT qual.forest_quartile,
       COUNT(qual.country_name)
FROM qual
GROUP BY 1
```

# Top Quartile Countries, 2016:
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. # 9 countries had a forestation percentage greater than 75% in 2016

```
WITH qual AS (
    SELECT country_name,
           region,
           ROUND(percent_of_the_land_area_as_forest::DECIMAL,2) AS
percent_forest,
           CASE WHEN percent_of_the_land_area_as_forest <=25 THEN '1'
                WHEN percent_of_the_land_area_as_forest <=50 THEN '2'
                WHEN percent_of_the_land_area_as_forest <=75 THEN '3'
                ELSE '4' END AS forest_quartile
    FROM forestation
    WHERE year = 2016
    AND forest_area_sqkm IS NOT NULL
    AND total_area_sqkm IS NOT NULL
    ORDER BY percent_of_the_land_area_as_forest)
```

```
SELECT qual.country_name,
       qual.region,
       qual.percent_forest
FROM qual
WHERE qual.forest_quartile = '4'
ORDER BY 3 DESC;
```