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 km2** in 1990. As of 2016, the most recent year for which data was available, that number had fallen to **39958245.9 km2,** a loss of **1324449 km2**, or **3.208**%.

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** km2).

## 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 |  |
| Latin America & Caribbean | 51.03 | 46.16 |  |
| Europe & Central Asia | 37.27 | 38.04 |  |
| North America | 35.65 | 36.04 |  |
| Sub-Saharan Africa | 30.67 | 28.79 |  |
| East Asia & Pacific | 25.77 | 26.36 |  |
| South Asia | 16.51 | 17.51 |  |
| Middle East & North Africa | 1.78 | 2.07 |  |
| 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**

### 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 **527229.06** **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 **79200**, **km2** 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.

### 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 |
| Indonesia | East Asia & Pacific | 282194 |
| Myanmar | East Asia & Pacific | 107234 |
| Nigeria | Sub-Saharan Africa | 106506 |
| Tanzania | Sub-Saharan Africa | 102320 |

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 |
| Mauritiana | 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 Mauritiana**. 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.

### QUARTILES

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

|  |  |
| --- | --- |
| Quartile | Number of Countries |
| 1 | 85 |
| 2 | 72 |
| 3 | 38 |
| 4 | 9 |

The largest number of countries in 2016 were found in the **First** 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?*
* *Which countries should we focus on over others?*

**Global Situation:**

From 1990 to 2016, the world has lost over **1324449 km2** of forest land, which amounts to 3.08% decrease since 1990. In order to understand the size of this forest land lost, this area is slightly less than the area of Peru.

**Regional Outlook:**

From Table 2.1 we see that Latin America & Caribbean and Sub-Saharan Africa have seen a drop of 4.87 and 4.63 % repectively in the amount of forest land during the period of 1990-2016.

Howeverr other regions have shown and increase on the percentage forest land. But, overall the world has seen a drop of 1.04% in forest land

**Country Outlook:**

We see that China and US have seen an increase in the total forest area by **527229.06 km2** and **79200 km2**. Since, these countries are very large in area these numbers would be better represented in terms of percentage increase.

From Table 3.1 we obser that Brazil and Indonesia have large scale deforestation followed by Myanmar.

From Table 3.2 we observe that Sub-Saharan region countries have experience large percentage forest change. The number are mind-boggeling, with Togo loosing almost 75% followed by Nigeria, Uganda, Mauritiana and Honduras loosing approximately 62%, 59%, 47%, and 45% respectively.

From Table 3.3 we observe that only 9 countries have a percentage forest land greater than 75%, while 85 countries have a forest land of less than 25%.

Table 3.4 list the name of countries having the largest percentage of forest area.

**Suggestions:**

**1.** Countries like Brazil, indonesia, Myanmar should be focussed give they are loosing so much forest land

2. regions of Latin America & Caribbean Sub-Saharan Africa need more focus

3. Nigeria needs special efforts given they are in top 5 bothe interms of % forest land lost and total forest land lost.

4. We also focus on region where the forest area has increased and take efforts to preserve it and understand what steps these countries take.

Github: [https://github.com/rishabhCMS/SQL\_Deforestation\_project#sql-deforestation-project](https://github.com/rishabhCMS/SQL_Deforestation_project" \l "sql-deforestation-project)

**SQL Queries**

**create view**

CREATE VIEW forestation AS

SELECT f.country\_code,

f.country\_name,

r.region,

r.income\_group,

f.year,

f.forest\_area\_sqkm,

l.total\_area\_sq\_mi\*2.59 AS land\_area\_sqkm,

f.forest\_area\_sqkm/(l.total\_area\_sq\_mi\*2.59)\*100 AS prcnt\_area

FROM forest\_area f

INNER JOIN land\_area l

ON f.country\_code = l.country\_code

AND f.year = l.year

JOIN regions r

ON r.country\_code = f.country\_code

Order by 1,5

## Part 1: Global Situation

**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 \* FROM forest\_area

WHERE country\_name= 'World' AND year = '1990'

**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 \* FROM forest\_area

WHERE country\_name= 'World' AND year = '2016'

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

WITH t1990 AS (SELECT forest\_area\_sqkm

FROM forest\_area

WHERE country\_name = 'World' AND year = '1990'),

t2016 AS (SELECT forest\_area\_sqkm

FROM forest\_area

WHERE country\_name = 'World' AND year = '2016')

SELECT (t1990.forest\_area\_sqkm - t2016.forest\_area\_sqkm) change

FROM t1990, t2016

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

WITH t1990 AS (SELECT forest\_area\_sqkm

FROM forest\_area

WHERE country\_name = 'World' AND year = '1990'),

t2016 AS (SELECT forest\_area\_sqkm

FROM forest\_area

WHERE country\_name = 'World' AND year = '2016')

SELECT t1990.forest\_area\_sqkm forest\_area\_sqkm\_1990,

t2016.forest\_area\_sqkm forest\_area\_sqkm\_2016,

(t1990.forest\_area\_sqkm - t2016.forest\_area\_sqkm) change,

((t1990.forest\_area\_sqkm - t2016.forest\_area\_sqkm)/t1990.forest\_area\_sqkm)\*100 as prcnt\_change

FROM t1990, t2016

**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?**

WITH Country\_2016 AS (SELECT \*

FROM land\_area

WHERE year = '2016')

SELECT \*,

ABS((((Country\_2016.total\_area\_sq\_mi)\*2.59) - 1324449)) diff\_sq\_km

FROM Country\_2016

ORDER BY diff\_sq\_km

LIMIT 10

**Part 2a.**

**I. What was the percent forest of the entire world in 2016?**

WITH land\_table AS (SELECT \*

FROM land\_area

WHERE year = '2016' AND country\_name = 'World'),

forest\_table AS (SELECT \*

FROM forest\_area

WHERE year = '2016' AND country\_name = 'World')

SELECT r.region,

SUM(l.total\_area\_sq\_mi\*2.59) total\_area\_sqkm,

SUM(f.forest\_area\_sqkm) total\_forest\_area\_sqkm,

ROUND(

cast((SUM(f.forest\_area\_sqkm)/(SUM(l.total\_area\_sq\_mi\*2.59)))\*100

AS NUMERIC),2

) AS prcnt\_area

FROM land\_table l

INNER JOIN forest\_table f

ON f.country\_name = l.country\_name

INNER JOIN regions r

ON r.country\_code = f.country\_code

GROUP BY 1

ORDER BY 4 DESC

LIMIT 10

**ii. Which region had the HIGHEST percent forest in 2016**

WITH land\_table AS (SELECT \*

FROM land\_area

WHERE year = '2016'),

forest\_table AS (SELECT \*

FROM forest\_area

WHERE year = '2016')

SELECT r.region,

SUM(l.total\_area\_sq\_mi\*2.59) total\_area\_sqkm,

SUM(f.forest\_area\_sqkm) total\_forest\_area\_sqkm,

ROUND(

cast((SUM(f.forest\_area\_sqkm)/(SUM(l.total\_area\_sq\_mi\*2.59)))\*100

AS NUMERIC),2

) AS prcnt\_area

FROM land\_table l

INNER JOIN forest\_table f

ON f.country\_name = l.country\_name

INNER JOIN regions r

ON r.country\_code = f.country\_code

GROUP BY 1

ORDER BY 4 DESC

LIMIT 10

**iii.  which had the LOWEST, to 2 decimal places?**

WITH land\_table AS (SELECT \*

FROM land\_area

WHERE year = '2016'),

forest\_table AS (SELECT \*

FROM forest\_area

WHERE year = '2016')

SELECT r.region,

SUM(l.total\_area\_sq\_mi\*2.59) total\_area\_sqkm,

SUM(f.forest\_area\_sqkm) total\_forest\_area\_sqkm,

ROUND(

cast((SUM(f.forest\_area\_sqkm)/(SUM(l.total\_area\_sq\_mi\*2.59)))\*100

AS NUMERIC),2

) AS prcnt\_area

FROM land\_table l

INNER JOIN forest\_table f

ON f.country\_name = l.country\_name

INNER JOIN regions r

ON r.country\_code = f.country\_code

GROUP BY 1

ORDER BY 4 ASC

LIMIT 10

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

WITH land\_table AS (SELECT \*

FROM land\_area

WHERE year = '1990' and country\_name = 'World' AND land\_area.total\_area\_sq\_mi IS NOT NULL),

forest\_table AS (SELECT \*

FROM forest\_area

WHERE year = '1990' and country\_name = 'World' AND forest\_area.forest\_area\_sqkm IS NOT NULL)

SELECT r.region,

SUM(l.total\_area\_sq\_mi\*2.59) total\_area\_sqkm,

SUM(f.forest\_area\_sqkm) total\_forest\_area\_sqkm,

ROUND(

cast((SUM(f.forest\_area\_sqkm)/(SUM(l.total\_area\_sq\_mi\*2.59)))\*100

AS NUMERIC),2

) AS prcnt\_area

FROM land\_table l

INNER JOIN forest\_table f

ON f.country\_name = l.country\_name

INNER JOIN regions r

ON r.country\_code = f.country\_code

GROUP BY 1

ORDER BY 4 DESC

LIMIT 10

**b ii) Which region had the HIGHEST percent forest in 1990?**

WITH land\_table AS (SELECT \*

FROM land\_area

WHERE year = '1990'),

forest\_table AS (SELECT \*

FROM forest\_area

WHERE year = '1990')

SELECT r.region,

SUM(l.total\_area\_sq\_mi\*2.59) total\_area\_sqkm,

SUM(f.forest\_area\_sqkm) total\_forest\_area\_sqkm,

ROUND(

cast((SUM(f.forest\_area\_sqkm)/(SUM(l.total\_area\_sq\_mi\*2.59)))\*100

AS NUMERIC),2

) AS prcnt\_area

FROM land\_table l

INNER JOIN forest\_table f

ON f.country\_name = l.country\_name

INNER JOIN regions r

ON r.country\_code = f.country\_code

GROUP BY 1

ORDER BY 4 DESC

LIMIT 10

**b iii. which had the LOWEST, to 2 decimal places?**

WITH land\_table AS (SELECT \*

FROM land\_area

WHERE year = '1990'),

forest\_table AS (SELECT \*

FROM forest\_area

WHERE year = '1990')

SELECT r.region,

SUM(l.total\_area\_sq\_mi\*2.59) total\_area\_sqkm,

SUM(f.forest\_area\_sqkm) total\_forest\_area\_sqkm,

ROUND(

cast((SUM(f.forest\_area\_sqkm)/(SUM(l.total\_area\_sq\_mi\*2.59)))\*100

AS NUMERIC),2

) AS prcnt\_area

FROM land\_table l

INNER JOIN forest\_table f

ON f.country\_name = l.country\_name

INNER JOIN regions r

ON r.country\_code = f.country\_code

GROUP BY 1

ORDER BY 4 ASC

LIMIT 10

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

WITH t1 AS(WITH land\_table AS (SELECT \*

FROM land\_area

WHERE year = '1990'),

forest\_table AS (SELECT \*

FROM forest\_area

WHERE year = '1990')

SELECT r.region,

SUM(l.total\_area\_sq\_mi\*2.59) total\_area\_sqkm,

SUM(f.forest\_area\_sqkm) total\_forest\_area\_sqkm,

ROUND(

cast((SUM(f.forest\_area\_sqkm)/(SUM(l.total\_area\_sq\_mi\*2.59)))\*100

AS NUMERIC),2

) AS prcnt\_area

FROM land\_table l

INNER JOIN forest\_table f

ON f.country\_name = l.country\_name

INNER JOIN regions r

ON r.country\_code = f.country\_code

GROUP BY 1

ORDER BY 4 DESC),

t2 AS (WITH land\_table AS (SELECT \*

FROM land\_area

WHERE year = '2016'),

forest\_table AS (SELECT \*

FROM forest\_area

WHERE year = '2016')

SELECT r.region,

SUM(l.total\_area\_sq\_mi\*2.59) total\_area\_sqkm,

SUM(f.forest\_area\_sqkm) total\_forest\_area\_sqkm,

ROUND(

cast((SUM(f.forest\_area\_sqkm)/(SUM(l.total\_area\_sq\_mi\*2.59)))\*100

AS NUMERIC),2

) AS prcnt\_area

FROM land\_table l

INNER JOIN forest\_table f

ON f.country\_name = l.country\_name

INNER JOIN regions r

ON r.country\_code = f.country\_code

GROUP BY 1

ORDER BY 4 DESC)

SELECT t1.region, t1.prcnt\_area prcnt\_area\_1990, t2.prcnt\_area prcnt\_area\_2016,

t1.prcnt\_area - t2.prcnt\_area diff

FROM t1

JOIN t2

ON t1.region = t2.region

ORDER BY 1 DESC

**Part 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?**

WITH t1990 AS (SELECT \*

FROM forest\_area

WHERE country\_name != 'World' AND year = '1990' AND forest\_area\_sqkm IS NOT NULL),

t2016 AS (SELECT \*

FROM forest\_area

WHERE country\_name != 'World' AND year = '2016' AND forest\_area\_sqkm IS NOT NULL)

SELECT t1990.country\_name,

t1990.forest\_area\_sqkm forest\_area\_sqkm\_1990,

t2016.forest\_area\_sqkm forest\_area\_sqkm\_2016,

(t1990.forest\_area\_sqkm - t2016.forest\_area\_sqkm) change,

ROUND(

CAST(((t1990.forest\_area\_sqkm - t2016.forest\_area\_sqkm)/t1990.forest\_area\_sqkm)\*100 AS NUMERIC),2

) as prcnt\_change

FROM t1990

JOIN t2016

ON t1990.country\_name = t2016.country\_name

ORDER BY 4 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 t1990 AS (SELECT \*

FROM forest\_area

WHERE country\_name != 'World' AND year = '1990' AND forest\_area\_sqkm IS NOT NULL),

t2016 AS (SELECT \*

FROM forest\_area

WHERE country\_name != 'World' AND year = '2016' AND forest\_area\_sqkm IS NOT NULL)

SELECT t1990.country\_name,

t1990.forest\_area\_sqkm forest\_area\_sqkm\_1990,

t2016.forest\_area\_sqkm forest\_area\_sqkm\_2016,

(t1990.forest\_area\_sqkm - t2016.forest\_area\_sqkm) change,

ROUND(

CAST(((t1990.forest\_area\_sqkm - t2016.forest\_area\_sqkm)/t1990.forest\_area\_sqkm)\*100 AS NUMERIC),2

) as prcnt\_change

FROM t1990

JOIN t2016

ON t1990.country\_name = t2016.country\_name

ORDER BY 5 DESC

LIMIT 5

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

WITH land\_table AS (SELECT \*

FROM land\_area

WHERE year = '2016' and country\_name != 'World' AND total\_area\_sq\_mi IS NOT NULL),

forest\_table AS (SELECT \*

FROM forest\_area

WHERE year = '2016' and country\_name != 'World' AND forest\_area\_sqkm IS NOT NULL),

t1 AS (SELECT f.country\_name,

SUM(l.total\_area\_sq\_mi\*2.59) total\_area\_sqkm,

SUM(f.forest\_area\_sqkm) total\_forest\_area\_sqkm,

ROUND(

cast((SUM(f.forest\_area\_sqkm)/(SUM(l.total\_area\_sq\_mi\*2.59)))\*100

AS NUMERIC),2

) AS prcnt\_area

FROM land\_table l

INNER JOIN forest\_table f

ON f.country\_name = l.country\_name

INNER JOIN regions r

ON r.country\_code = f.country\_code

GROUP BY 1

ORDER BY 4 DESC

)

SELECT t2.quartile, COUNT(\*)

FROM(SELECT \*,

CASE

WHEN t1.prcnt\_area > 75 THEN 4

WHEN t1.prcnt\_area > 50 AND t1.prcnt\_area <= 75 THEN 3

WHEN t1.prcnt\_area > 25 AND t1.prcnt\_area <= 50 THEN 2

ELSE 1

END AS quartile

FROM t1) AS t2

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 land\_table AS (SELECT \*

FROM land\_area

WHERE year = '2016' and country\_name != 'World'),

forest\_table AS (SELECT \*

FROM forest\_area

WHERE year = '2016' and country\_name != 'World'),

t1 AS (SELECT f.country\_name,

SUM(l.total\_area\_sq\_mi\*2.59) total\_area\_sqkm,

SUM(f.forest\_area\_sqkm) total\_forest\_area\_sqkm,

ROUND(

cast((SUM(f.forest\_area\_sqkm)/(SUM(l.total\_area\_sq\_mi\*2.59)))\*100

AS NUMERIC),2

) AS prcnt\_area

FROM land\_table l

INNER JOIN forest\_table f

ON f.country\_name = l.country\_name

INNER JOIN regions r

ON r.country\_code = f.country\_code

GROUP BY 1

ORDER BY 4 DESC

)

SELECT \*, COUNT(\*)

FROM(SELECT \*,

CASE

WHEN t1.prcnt\_area > 75 THEN 4

WHEN t1.prcnt\_area > 50 AND t1.prcnt\_area <= 75 THEN 3

WHEN t1.prcnt\_area > 25 AND t1.prcnt\_area <= 50 THEN 2

ELSE 1

END AS quartile

FROM t1) AS t2

WHERE t2.quartile = 4

GROUP BY 1,2,3,4,5

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

WITH land\_table AS (SELECT \*

FROM land\_area

WHERE year = '2016' and country\_name != 'World'),

forest\_table AS (SELECT \*

FROM forest\_area

WHERE year = '2016' and country\_name != 'World'),

t1 AS (SELECT f.country\_name,

SUM(l.total\_area\_sq\_mi\*2.59) total\_area\_sqkm,

SUM(f.forest\_area\_sqkm) total\_forest\_area\_sqkm,

ROUND(

cast((SUM(f.forest\_area\_sqkm)/(SUM(l.total\_area\_sq\_mi\*2.59)))\*100

AS NUMERIC),2

) AS prcnt\_area

FROM land\_table l

INNER JOIN forest\_table f

ON f.country\_name = l.country\_name

INNER JOIN regions r

ON r.country\_code = f.country\_code

GROUP BY 1

ORDER BY 4 DESC

)

SELECT COUNT(\*)

FROM(SELECT \*,

CASE

WHEN t1.prcnt\_area > 75 THEN 4

WHEN t1.prcnt\_area > 50 AND t1.prcnt\_area <= 75 THEN 3

WHEN t1.prcnt\_area > 25 AND t1.prcnt\_area <= 50 THEN 2

ELSE 1

END AS quartile

FROM t1) As t2

WHERE t2.prcnt\_area > (

SELECT prcnt\_area

FROM t1

WHERE country\_name = 'United States'

)