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.

# GLOBAL SITUATION

According to the World Bank, the total forest area of the world was 899962748.20sqkm in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 8710897606.20sqkm, a loss of 288729882.00sqkm, 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 279039997.62sqkm.

# 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 |
| 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%.

# COUNTRY-LEVEL DETAIL

### SUCCESS STORIES

There is one particularly bright spot in the data at the country level, over time. This country actually increased in forest area from 1990 to 2016 by 527229.062. 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 United States, but it only saw an increase of 79200, much lower than the figure for China.

Russian Federation 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 | 5467050 |
| Indonesia | East Asia & Pacific | 1185450 |
| Myanmar | East Asia & Pacific | 392180 |
| Nigeria | Sub-Saharan Africa | 172340 |
| Tanzania | Sub-Saharan Africa | 559200 |

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 | 59.13 |
| 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.

### 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 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 |
| Seychelles | Sub\_Saharan Africa | 88.41 |
| Palau | East Asia & Pacific | 87.61 |
| American Samoa | East Asia & Pacific | 87.5 |
| Guyana | Latin America & Caribbean | 83.9 |
| Lao PDR | East Asia & Pacific | 82.11 |
| Solomon Islands | East Asia & Pacific | 77.86 |

## RECOMMENDATIONS

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

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

World Bank data shows how the forest area of the entire world was decreased by 3% whch is slightly more than the entire land area of Peru. The results from the data also shows us how the global situation, regional outlook and country level details of the data can help the organization to fulfill their mission of raising awareness about deforestation

and make an impact on different levels throughout the world. There are success stories as well as biggest concerns about how the data has changed over time.

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

94 countries had a percent forestation higher than United States in 2016. While there are bigger countries with lager forest area, focus should be given to those countries who lost forest area in larger amount and percent like Brail, Indonesia, Myanmar, Nigeria, Tongo, Uganda. Specifically, most of these countries fall under Sub-Saharan Africa region.

## APPENDIX: QUERIES

* CREATING VIEW:

DROP VIEW IF EXISTS forestation;

CREATE VIEW forestation AS (

SELECT

r.country\_name AS r\_country\_name, r.country\_code AS r\_country\_code, income\_group, region AS region, la.country\_code AS la\_country\_code, la.country\_name AS la\_country\_name, la.year AS la\_year,

(la.total\_area\_sq\_mi\*2.59) AS la\_total\_area\_sqkm, fa.country\_code AS fa\_country\_code, fa.country\_name AS fa\_country\_name,

fa.year AS fa\_year,

fa.forest\_area\_sqkm AS fa\_forest\_area\_sqkm, fa.forest\_area\_sqkm/(la.total\_area\_sq\_mi \*2.59)\*100 AS forest\_area\_percent FROM land\_area AS la

JOIN forest\_area AS fa ON fa.year = la.year JOIN regions AS r

ON r.country\_code = la.country\_code

);

* total forest area (in sq km) of the world in 1990 AND 2016

SELECT fa\_country\_name,

fa\_year,ROUND(SUM(fa\_forest\_area\_sqkm)::numeric,2) FROM forestation

GROUP BY 1,2

HAVING fa\_country\_name = 'World' AND fa\_year =1990 (2016)

* the change (in sq km) in the forest area of the world from 1990 to 2016
* the percent change in forest area of the world between 1990 and 2016

WITH forest\_2016 AS (SELECT fa\_country\_name,

fa\_year,ROUND(SUM(fa\_forest\_area\_sqkm)::numeric,2) AS area2016

FROM forestation GROUP BY 1,2

HAVING fa\_country\_name = 'World' AND fa\_year =2016),

forest\_1990 AS (SELECT fa\_country\_name, fa\_year,ROUND(SUM(fa\_forest\_area\_sqkm)::numeric,2) AS area1990

FROM forestation GROUP BY 1,2

HAVING fa\_country\_name = 'World' AND fa\_year =1990

)

SELECT f2016.area2016,

f1990.area1990,

f1990.area1990-f2016.area2016 AS difference, ROUND((f1990.area1990-f2016.area2016)/f1990.area1990 \*

100::numeric,2) AS f\_percent

FROM forest\_2016 AS f2016, forest\_1990 AS f1990

* 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 la\_year,

la\_country\_name, ROUND(SUM(la\_total\_area\_sqkm) ::numeric,2 )

FROM forestation GROUP BY 1,2

HAVING SUM(la\_total\_area\_sqkm) < 288729882 AND la\_year = 2016 ORDER BY SUM(la\_total\_area\_sqkm) DESC

LIMIT 1

* CREATE VIEW REG\_DIST

CREATE OR REPLACE VIEW reg\_dist AS

SELECT r.region, la.year,

SUM(fa.forest\_area\_sqkm) total\_forest\_area\_sqkm, SUM(la.total\_area\_sq\_mi\*2.59) AS total\_area\_sqkm, (SUM(fa.forest\_area\_sqkm)/SUM(la.total\_area\_sq\_mi\*2.59)) \* 100 AS

percent\_fa\_region

FROM forest\_area fa

JOIN land\_area la

ON fa.country\_code = la.country\_code AND fa.year = la.year JOIN regions r

ON la.country\_code = r.country\_code GROUP BY 1,2

ORDER BY 1,2

* The percent forest of the entire world in 2016

SELECT ROUND(CAST(percent\_fa\_region AS numeric),2) AS percent\_fa\_region FROM reg\_dist

WHERE year = 2016 AND region = 'World';

* Which region had the HIGHEST percent forest in 2016

SELECT region,

ROUND(CAST(total\_area\_sqkm AS NUMERIC),2) AS total\_area\_sqkm, ROUND(CAST(percent\_fa\_region AS NUMERIC),2) AS percent\_fa\_region FROM reg\_dist

WHERE ROUND(CAST(percent\_fa\_region AS NUMERIC),2) = (SELECT MAX(

ROUND(

CAST(percent\_fa\_region AS

numeric),2

AND year=2016;

)

) AS max\_percent

WHERE year = 2016

)

FROM reg\_dist

* which had the LOWEST, to 2 decimal places?

SELECT region,

ROUND(CAST(total\_area\_sqkm AS NUMERIC),2) AS total\_area\_sqkm, ROUND(CAST(percent\_fa\_region AS NUMERIC),2) AS percent\_fa\_region FROM reg\_dist

WHERE ROUND(CAST(percent\_fa\_region AS NUMERIC),2) = (SELECT MIN(

ROUND(

CAST(percent\_fa\_region AS numeric),2

)

) AS min\_percent

AND year = 2016;

WHERE year = 2016

)

FROM reg\_dist

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

SELECT ROUND(CAST(percent\_fa\_region AS numeric),2) AS percent\_fa\_region FROM reg\_dist

WHERE year = 1990 AND region = 'World';

* Which region had the HIGHEST percent forest in 1990, and

SELECT region,

ROUND(CAST(total\_area\_sqkm AS NUMERIC),2) AS total\_area\_sqkm, ROUND(CAST(percent\_fa\_region AS NUMERIC),2) AS percent\_fa\_region FROM reg\_dist

WHERE ROUND(CAST(percent\_fa\_region AS NUMERIC),2) = (SELECT MAX(

ROUND(

CAST(percent\_fa\_region AS

numeric),2

AND year=1990;

)

) AS max\_percent

WHERE year = 1990

)

FROM reg\_dist

* which had the LOWEST, to 2 decimal places?

SELECT region,

ROUND(CAST(total\_area\_sqkm AS NUMERIC),2) AS total\_area\_sqkm, ROUND(CAST(percent\_fa\_region AS NUMERIC),2) AS percent\_fa\_region FROM reg\_dist

WHERE ROUND(CAST(percent\_fa\_region AS NUMERIC),2) = (SELECT MIN(

ROUND(

CAST(percent\_fa\_region AS numeric),2

)

) AS min\_percent

AND year = 1990;

WHERE year = 1990

)

FROM reg\_dist

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

WITH table1990 AS (SELECT \* FROM reg\_dist WHERE year =1990), table2016 AS (SELECT \* FROM reg\_dist 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;

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

WITH table1990 AS (SELECT fa.country\_code,

fa.country\_name, fa.year, fa.forest\_area\_sqkm

FROM forest\_area fa

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

),

table2016 AS (SELECT fa.country\_code,

fa.country\_name, fa.year, fa.forest\_area\_sqkm

FROM forest\_area fa

WHERE fa.year = 2016 AND fa.forest\_area\_sqkm IS NOT NULL AND fa.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,

( table2016.forest\_area\_sqkm- table1990.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;

* Which 5 countries saw the largest percent increase in forest area from 1990 to 2016? What was the percent change to 2 decimal places for each?

WITH table1990 AS (SELECT fa.country\_code,

fa.country\_name, fa.year, fa.forest\_area\_sqkm

FROM forest\_area fa

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

),

table2016 AS (SELECT fa.country\_code,

fa.country\_name, fa.year, fa.forest\_area\_sqkm

FROM forest\_area fa

WHERE fa.year = 2016 AND fa.forest\_area\_sqkm IS NOT NULL AND fa.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.forest\_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\_area\_sqkm\*100) AS NUMERIC),2) DESC

LIMIT 5;

* Top 5 Amount Decrease in Forest Area by Country, 1990 & 2016:

WITH table1990 AS (SELECT fa.country\_code,

fa.country\_name, fa.year, fa.forest\_area\_sqkm

FROM forest\_area fa

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

),

table2016 AS (SELECT fa.country\_code,

fa.country\_name, fa.year, fa.forest\_area\_sqkm

FROM forest\_area fa

WHERE fa.year = 2016 AND fa.forest\_area\_sqkm IS NOT NULL AND fa.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,

ABS(table1990.forest\_area\_sqkm-table2016.forest\_area\_sqkm) AS diff\_fa\_sqkm,

ABS(ROUND(CAST(((table2016.forest\_area\_sqkm- table1990.forest\_area\_sqkm)/table1990.forest\_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 table1990.forest\_area\_sqkm-table2016.forest\_area\_sqkm DESC LIMIT 5;

* Top 5 percentDecrease in Forest Area by Country, 1990 & 2016:

WITH table1990 AS (SELECT fa.country\_code,

fa.country\_name, fa.year, fa.forest\_area\_sqkm

FROM forest\_area fa

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

),

table2016 AS (SELECT fa.country\_code,

fa.country\_name, fa.year, fa.forest\_area\_sqkm

FROM forest\_area fa

WHERE fa.year = 2016 AND fa.forest\_area\_sqkm IS NOT NULL AND fa.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.forest\_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\_area\_sqkm\*100) AS NUMERIC),2)

LIMIT 5;

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

With T1 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,

perc\_fa

(fa.forest\_area\_sqkm/(la.total\_area\_sq\_mi\*2.59))\*100 AS

FROM forest\_area fa JOIN land\_area la

ON fa.country\_code = la.country\_code

AND (fa.country\_name != 'World' AND fa.forest\_area\_sqkm IS

NOT NULL AND la.total\_area\_sq\_mi IS NOT NULL) AND (fa.year=2016 AND la.year = 2016) ORDER BY 6 DESC

),

t2 AS (SELECT t1.country\_code,

t1.country\_name, t1.year, t1.perc\_fa,

CASE WHEN t1.perc\_fa >= 75 THEN 4

WHEN t1.perc\_fa < 75 AND t1.perc\_fa >= 50 THEN 3 WHEN t1.perc\_fa < 50 AND t1.perc\_fa >=25 THEN 2

ELSE 1

END AS quartile

FROM t1 ORDER BY 5 DESC

)

SELECT t2.quartile ,

COUNT(t2.quartile) AS number\_of\_countries FROM t2

GROUP BY 1

ORDER BY 2 DESC;

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

With T1 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,

(fa.forest\_area\_sqkm/(la.total\_area\_sq\_mi\*2.59))\*100 AS

perc\_fa

FROM forest\_area fa JOIN land\_area la

ON fa.country\_code = la.country\_code

AND (fa.country\_name != 'World' AND fa.forest\_area\_sqkm IS

NOT NULL AND la.total\_area\_sq\_mi IS NOT NULL) AND (fa.year=2016 AND la.year = 2016) ORDER BY 6 DESC

),

t2 AS (SELECT t1.country\_code,

t1.country\_name, t1.year, t1.perc\_fa,

CASE WHEN t1.perc\_fa >= 75 THEN 4

WHEN t1.perc\_fa < 75 AND t1.perc\_fa >= 50 THEN 3 WHEN t1.perc\_fa < 50 AND t1.perc\_fa >=25 THEN 2

ELSE 1

END AS percentile

FROM t1 ORDER BY 5 DESC

)

SELECT t2.country\_name,r.region,ROUND(CAST(t2.perc\_fa AS NUMERIC),2 ) AS perc\_fa, t2.percentile

FROM t2

JOIN regions r

ON t2.country\_code = r.country\_code WHERE t2.percentile = 4

ORDER BY 3 DESC

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

With t1 AS (SELECT fa.country\_code,

fa.country\_name,

perc\_fa

fa.year, fa.forest\_area\_sqkm,

la.total\_area\_sq\_mi\*2.59 AS total\_area\_sqkm, (fa.forest\_area\_sqkm/(la.total\_area\_sq\_mi\*2.59))\*100 AS

FROM forest\_area fa JOIN land\_area la

ON fa.country\_code = la.country\_code

AND (fa.country\_name != 'World' AND fa.forest\_area\_sqkm IS

NOT NULL AND la.total\_area\_sq\_mi IS NOT NULL) AND (fa.year=2016 AND la.year = 2016) ORDER BY 6 DESC

)

SELECT COUNT(t1.country\_name) FROM t1

WHERE t1.perc\_fa > (SELECT t1.perc\_fa

FROM t1

WHERE t1.country\_name = 'United States'

)