Report for ForestQuery into GlobalDeforestation (1990 - 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

Great job all your answers in this section are correct, Kudos.

According to the World Bank, the total forest area of the world was **41282694.90 km^2** in 1990. As of 2016, the most recent year for which data was available, that number had fallen to **39958245.90 km^2**, a loss of **1324449 km^2** 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 **1279999.99 km^2**).

Yes, **Peru i**s the country with the total land area closest to the amount of forest area lost between 1990 and 2016

2. REGIONAL OUTLOOK

The regional outlook section contains the correct answers

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 **527229.062 km^2**. 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 km^2**, 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.

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

All the values in this table are correct, Kudos!

Country	Region	Absolute Forest Area Change
Brazil	Latin America & Caribbean	-541510
Indonesia	East Asia & Pacific	-282193.9844
Myanmar	East Asia & Pacific	-107234.0039
Nigeria	Sub-Saharan Africa	-106506.001
Tanzania	Sub-Saharan Africa	-102320

(*NOTICE*: Sign **'-'** (*minus*) in the last column means ***decrease***)

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

All the values in this table are correct, Kudos!

Country	Region	Pct Forest Area Change
Togo	Sub-Saharan Africa	-75.44
Nigeria	Sub-Saharan Africa	-61.80
Uganda	Sub-Saharan Africa	-59.27
Mauritania	Sub-Saharan Africa	-46.75
Honduras	Latin America & Caribbean	-45.03

(*NOTICE*: Sign **'-'** (*minus*) in the last column means ***decrease***)

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

The number of countries in the **second** quartile is slightly off. You should filter out the country value denoted as **"World"** in the WHERE clause for this query because it isn't a country.

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

Quartiles	Number of Countries
0 - 25%	85
25% - 50%	73
50% - 75%	38
75% - 100%	9

The largest number of countries in 2016 were found in the **1st (or '0 – 25%')** 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
American Samoa	East Asia & Pacific	87.50
Gabon	Sub-Saharan Africa	90.04
Guyana	Latin America & Caribbean	83.90
Lao PDR	East Asia & Pacific	82.11
Micronesia, Fed. Sts.	East Asia & Pacific	91.86
Palau	East Asia & Pacific	87.61
Seychelles	Sub-Saharan Africa	88.41
Solomon Islands	East Asia & Pacific	77.86
Suriname	Latin America & Caribbean	98.26

There are **94 countries** with their percent forestation higher than that of the United States.

4. RECOMMENDATIONS

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

1. What have you learned from the World Bank data?

- The result of deforestation is astonishing. In just 26 years from 1990 to 2016, the world's total forest area has dropped by *3.208%* or **1324449 km^2**, which is about the land size of **Peru**.
- Although there are only 2 out of 7 regions seeing dropping in forest coverage, the general trend for world's forestation area is decreasing. Especially **Latin America & Caribbean,** being the region with the largest forest coverage ever, has dropped nearly **5%** from **1990 (51.03%)** to **2016 (46.16%)**.
- **China** is on the top the increasing list in terms of forestation area from 1990 to 2016, over 6 times as large as that of the **United States**, which holds the 2nd place. There must be something interesting to explore and study.
- **4** out of the t**op 5** countries whose forest area decreased the most from 1990 to 2016 are in the **Sub-Saharan Africa** region. Specifically, the country of **Nigeria** is both referred to in either absolute forest area decrease or percentage decrease.
- In the **218** countries being investigated, there are **85** countries whose forest coverage rate is less than **25%**, the largest number of countries when grouped in forestation percent quartile. Only **47** countries have their forest coverage percentage above **75%**.

2. Which countries should we focus on over others?

Togo, Nigeria, Uganda and Mauritania in the **Sub-Saharan Africa** and **Honduras** need more attention as they are the countries that decreased in percentage of forest area the most. We need further exploration and study on what caused these decreasing and what actions could be taken, such as in law, education, economic aid, resource allocation, population control etc.

 In general speaking, deforestation is becoming more and more serious in our modern world. We should make deliberate plans and take instant action against it. If not, a sequence of chain events could happen such as climate change, air pollution, wildlife extinction and so on.

5. APPENDIX: SQL Queries Used

```sql
1   -- Create the VIEW
2   DROP VIEW IF EXISTS forestation;
3   CREATE VIEW forestation AS
4     SELECT f.country_code country_code,
5            f.country_name country_name,
6            f.year,
7            f.forest_area_sqkm,
8            l.total_area_sq_mi*2.59 total_area_sq_km,
9            100*f.forest_area_sqkm/(l.total_area_sq_mi*2.59)
    pct_forestation,
10           r.region,
11           r.income_group
12    FROM forest_area f
13    JOIN land_area l
14      ON f.country_code=l.country_code and f.year=l.year
15    JOIN regions r
16      ON r.country_code=f.country_code;
17
18  SELECT count(*)
19  FROM forestation;
20
21  -- PART I. GLOBAL SITUATION
22  -- Difference and percentage drop in forestation area betweeb 1990
    and 2016
23  WITH t1 AS (
24    SELECT *
25    FROM forestation
26    WHERE country_code='WLD' and year in (1990, 2016)
27    ORDER BY year)
28
29  SELECT year,
30    forest_area_sqkm,
31    LEAD(forest_area_sqkm) OVER (order by year) AS lead,
```

This creates a View successfully.

This produce the correct forest area values in 2016 and 1990, the forest loss and the percent change between 1990 and 2016.

```
32    LEAD(forest_area_sqkm) OVER (order by year)-forest_area_sqkm AS
   abs_diff,
33    ROUND((100*(LEAD(forest_area_sqkm) OVER (order by year)-
   forest_area_sqkm)/forest_area_sqkm)::NUMERIC, 3) AS pct_diff
34  FROM t1
35
36  -- Find the country with its land area in 2016 closest to the
   deforestation area between 1990 and 2016
37  WITH t1 AS (
38    SELECT *
39    FROM forestation
40    WHERE country_code='WLD' and year in (1990, 2016)
41    ORDER BY year),
42    t2 AS (
43      SELECT year,
44        forest_area_sqkm,
45        LEAD(forest_area_sqkm) OVER (order by year) AS lead,
46        LEAD(forest_area_sqkm) OVER (order by year)-forest_area_sqkm AS
   abs_diff,
47        ROUND((100*(LEAD(forest_area_sqkm) OVER (order by year)-
   forest_area_sqkm)/forest_area_sqkm)::NUMERIC, 3) AS pct_diff
48      FROM t1)
49
50  SELECT DISTINCT country_name,
51                  total_area_sq_km,
52                  (SELECT ABS(t2.abs_diff) FROM t2 ORDER BY year LIMIT
   1) AS abs_diff,
53                  ABS(total_area_sq_km-(SELECT ABS(t2.abs_diff) FROM t2
   ORDER BY year LIMIT 1)) AS diff
54  FROM forestation
55  ORDER BY 4
56
57  -- Part II. Regional Outlook
58  -- Find the world's forestation area percentage in 2016
59  SELECT country_code,
60         country_name,
61         year,
62         forest_area_sqkm,
63         total_area_sq_km,
64         ROUND(pct_forestation::NUMERIC, 2) pct_forestation
65  FROM forestation
66  WHERE year=2016 AND country_code='WLD';
67
68  -- Find the region with the highest forestation percentage in 2016
69  SELECT year,
70         region,
71         SUM(forest_area_sqkm) total_forestation,
```

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

> This correctly returns the percent of the forest area of the entire **World** in 2016 as **31.38**

```sql
72          SUM(total_area_sq_km) total_land,

    ROUND((100*SUM(forest_area_sqkm)/SUM(total_area_sq_km))::NUMERIC,2)
    forestation_pct
74  FROM forestation
75  GROUP BY 1,2
76  HAVING year=2016
77  ORDER BY forestation_pct DESC
78  LIMIT 1;

80  -- Find the region with the lowest forestation percentage in 2016
81  SELECT year,
82         region,
83         SUM(forest_area_sqkm) total_forestation,
84         SUM(total_area_sq_km) total_land,

    ROUND((100*SUM(forest_area_sqkm)/SUM(total_area_sq_km))::NUMERIC,2)
    forestation_pct
86  FROM forestation
87  GROUP BY 1,2
88  HAVING year=2016
89  ORDER BY forestation_pct
90  LIMIT 1;

92  -- Find the world's forestation area percentage in 1990
93  SELECT country_code,
94         country_name,
95         year,
96         forest_area_sqkm,
97         total_area_sq_km,
98         ROUND(pct_forestation::NUMERIC, 2) pct_forestation
99  FROM forestation
100 WHERE year=1990 AND country_code='WLD';

102 -- Find the region with the highest forestation percentage in 1990
103 SELECT year,
104        region,
105        SUM(forest_area_sqkm) total_forestation,
106        SUM(total_area_sq_km) total_land,

    ROUND((100*SUM(forest_area_sqkm)/SUM(total_area_sq_km))::NUMERIC,2)
    forestation_pct
108 FROM forestation
109 GROUP BY 1,2
110 HAVING year=1990
111 ORDER BY forestation_pct DESC
112 LIMIT 1;
```

> This produces **Latin America & Caribbean** as the region with the highest forest percentage in 2016

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

> This correctly returns the percent of the forest area of the entire **World** in 1990 as **32.42**

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

```sql
113
114  -- Find the region with the lowest forestation percentage in 1990
115  SELECT year,
116         region,
117           SUM(forest_area_sqkm) total_forestation,
118           SUM(total_area_sq_km) total_land,
119
     ROUND((100*SUM(forest_area_sqkm)/SUM(total_area_sq_km))::NUMERIC,2)
     forestation_pct
120  FROM forestation
121  GROUP BY 1,2
122  HAVING year=1990
123  ORDER BY forestation_pct
124  LIMIT 1;
125
126  -- Calculate Table 2.1: Percent Forest Area by Region, 1990 & 2016
127  DROP VIEW IF EXISTS t1;
128  CREATE VIEW t1 AS (
129  SELECT year yr,
130         region,
131           SUM(forest_area_sqkm) total_forestation,
132           SUM(total_area_sq_km) total_land,
133
     ROUND((100*SUM(forest_area_sqkm)/SUM(total_area_sq_km))::NUMERIC,2)
     forestation_pct
134  FROM forestation
135  GROUP BY 1,2
136  HAVING year in (1990, 2016)
137  ORDER BY region, yr);
138
139  WITH tab1 AS (
140      SELECT region,
141      forestation_pct
142      FROM t1
143      where yr=1990),
144
145      tab2 AS (
146        SELECT region,
147        forestation_pct
148        FROM t1
149        where yr=2016)
150
151  SELECT tab1.region,
152  tab1.forestation_pct AS pct_1990,
153  tab2.forestation_pct AS pct_2019
154  FROM tab1
155  JOIN tab2 ON tab1.region=tab2.region
```

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

> This list the various regions and their respective forest area percentages in 1990 and 2016

```sql
-- Part III. Country-level Detail
-- A. Success Stories
-- Largest change in terms of forest_area
WITH tab_1990 AS (
  SELECT country_code,
         country_name,
         forest_area_sqkm,
         total_area_sq_km,
         pct_forestation
  FROM forestation
  WHERE year=1990
  ORDER BY country_name),

  tab_2016 AS (
    SELECT country_code,
           country_name,
           forest_area_sqkm,
           total_area_sq_km,
           pct_forestation
    FROM forestation
    WHERE year=2016
    ORDER BY country_name),

  tab_join AS (
    SELECT tab_1990.country_name,
           tab_1990.forest_area_sqkm forest_1990,
           tab_2016.forest_area_sqkm forest_2016,
           tab_1990.total_area_sq_km land_1990,
           tab_2016.total_area_sq_km land_2016,
           tab_1990.pct_forestation pct_1990,
           tab_2016.pct_forestation pct_2016
    FROM tab_1990
    JOIN tab_2016 ON tab_1990.country_code=tab_2016.country_code)

SELECT country_name,
       forest_1990,
       forest_2016,
       (forest_2016-forest_1990) AS forest_area_change,
       (100*(pct_2016-pct_1990)/pct_1990) AS pct_change,
       land_1990,
       land_2016
FROM tab_join
WHERE forest_1990 IS NOT NULL AND forest_2016 IS NOT NULL AND
country_name!='World'
ORDER BY forest_area_change DESC
```

This correctly produces **China** and the **United States** as the top two countries in terms of forest area increase in terms of sqkm

```sql
-- Largest change in terms of forest_area percentage
WITH tab_1990 AS (
  SELECT country_code,
         country_name,
         forest_area_sqkm,
         total_area_sq_km,
         pct_forestation
  FROM forestation
  WHERE year=1990
  ORDER BY country_name),

  tab_2016 AS (
    SELECT country_code,
           country_name,
           forest_area_sqkm,
           total_area_sq_km,
           pct_forestation
    FROM forestation
    WHERE year=2016
    ORDER BY country_name),

  tab_join AS (
    SELECT tab_1990.country_name,
           tab_1990.forest_area_sqkm forest_1990,
           tab_2016.forest_area_sqkm forest_2016,
           tab_1990.total_area_sq_km land_1990,
           tab_2016.total_area_sq_km land_2016,
           tab_1990.pct_forestation pct_1990,
           tab_2016.pct_forestation pct_2016
    FROM tab_1990
    JOIN tab_2016 ON tab_1990.country_code=tab_2016.country_code)

SELECT country_name,
       forest_1990,
       forest_2016,
       (forest_2016-forest_1990) AS forest_area_change,
       (100*(pct_2016-pct_1990)/pct_1990) AS pct_change,
       land_1990,
       land_2016
FROM tab_join
WHERE forest_1990 IS NOT NULL AND forest_2016 IS NOT NULL AND
country_name!='World'
ORDER BY pct_change DESC

-- B. Largest Concerns
-- Calculate Table 3.1 Top 5 Amount Decrease in Forest Area by
Country, 1990 & 2016
```

> This correctly produces **Iceland** as the top country in terms of percent change increase in forest area

```sql
WITH tab_1990 AS (
  SELECT country_code,
         country_name,
         region,
         forest_area_sqkm,
         total_area_sq_km,
         pct_forestation
  FROM forestation
  WHERE year=1990
  ORDER BY country_name),

  tab_2016 AS (
    SELECT country_code,
           country_name,
           region,
           forest_area_sqkm,
           total_area_sq_km,
           pct_forestation
    FROM forestation
    WHERE year=2016
    ORDER BY country_name),

  tab_join AS (
    SELECT tab_1990.country_name,
           tab_1990.region,
           tab_1990.forest_area_sqkm forest_1990,
           tab_2016.forest_area_sqkm forest_2016,
           tab_1990.total_area_sq_km land_1990,
           tab_2016.total_area_sq_km land_2016,
           tab_1990.pct_forestation pct_1990,
           tab_2016.pct_forestation pct_2016
    FROM tab_1990
    JOIN tab_2016 ON tab_1990.country_code=tab_2016.country_code)

SELECT country_name,
       region,
       forest_1990,
       forest_2016,
       (forest_2016-forest_1990) AS forest_area_change,
       (100*(pct_2016-pct_1990)/pct_1990) AS pct_change,
       land_1990,
       land_2016
FROM tab_join
WHERE forest_1990 IS NOT NULL AND forest_2016 IS NOT NULL AND
country_name!='World'
ORDER BY forest_area_change
```

> Awesome, the top five countries in terms of sqkm forestation decrease are all correct. Brazil, Indonesia, Myanmar, Nigeria, and Tanzania

```sql
-- Calculate Table 3.2 Top 5 Percent Decrease in Forest Area by
Country, 1990 & 2016
WITH tab_1990 AS (
  SELECT country_code,
         country_name,
         region,
         forest_area_sqkm,
         total_area_sq_km,
         pct_forestation
  FROM forestation
  WHERE year=1990
  ORDER BY country_name),

  tab_2016 AS (
    SELECT country_code,
           country_name,
           region,
           forest_area_sqkm,
           total_area_sq_km,
           pct_forestation
    FROM forestation
    WHERE year=2016
    ORDER BY country_name),

  tab_join AS (
    SELECT tab_1990.country_name,
           tab_1990.region,
           tab_1990.forest_area_sqkm forest_1990,
           tab_2016.forest_area_sqkm forest_2016,
           tab_1990.total_area_sq_km land_1990,
           tab_2016.total_area_sq_km land_2016,
           tab_1990.pct_forestation pct_1990,
           tab_2016.pct_forestation pct_2016
    FROM tab_1990
    JOIN tab_2016 ON tab_1990.country_code=tab_2016.country_code)

SELECT country_name,
       region,
       forest_1990,
       forest_2016,
       (forest_2016-forest_1990) AS forest_area_change,
       ROUND((100*(pct_2016-pct_1990)/pct_1990)::NUMERIC, 2) AS
pct_change,
       land_1990,
       land_2016
FROM tab_join
```

Awesome, the top five countries in terms of percent decrease in forestation are all correct. **Togo, Nigeria, Uganda, Mauritania, Honduras**

```sql
337  WHERE forest_1990 IS NOT NULL AND forest_2016 IS NOT NULL AND
     country_name!='World'
338  ORDER BY pct_change
339
340  -- C. Quartiles
341  -- Calculate Table 3.3 Count of Countries Grouped by Forestation
     Percent Quartiles, 2016
342  WITH tab_quartile AS (
343    SELECT country_name,
344           pct_forestation
345    FROM forestation
346    WHERE year=2016 AND pct_forestation IS NOT NULL AND country_name<>'World'
347    ORDER BY 2),
348
349    tab_quartile1 AS (
350      SELECT country_name,
351             pct_forestation,
352             CASE
353               WHEN pct_forestation<=25 THEN '0 - 25%'
354               WHEN pct_forestation<=50 THEN '25% - 50%'
355               WHEN pct_forestation<=75 THEN '50% - 75%'
356               ELSE '75% - 100%'
357             END AS quartiles
358      FROM tab_quartile)
359
360  SELECT quartiles, count(country_name) number_of_countries
361  FROM tab_quartile1
362  GROUP BY 1
363  ORDER BY 1
364
365  -- List all of the countries that were in the 4th quartile (percent
     forest > 75%) in 2016.
366  WITH tab_quartile AS (
367    SELECT country_name,
368           region,
369           pct_forestation
370    FROM forestation
371    WHERE year=2016 AND pct_forestation IS NOT NULL
372    ORDER BY 2),
373
374    tab_quartile1 AS (
375      SELECT country_name,
376             region,
377             pct_forestation,
378             CASE
379               WHEN pct_forestation<=25 THEN '0 - 25%'
380               WHEN pct_forestation<=50 THEN '25% - 50%'
```

You should filter out the country value denoted as **"World"** from this query. *(annotation on lines 337–339)*

AND country_name<>'World' *(annotation on line 346)*

Yes, 9 countries had a forestation percentage greater than 75% in 2016 *(annotation near line 366)*

You could get this same output with the query below
**SELECT country_name, region, pct_forestation
FROM forestation
WHERE pct_forestation>75 AND year=2016** *(annotation near line 372)*

```sql
                WHEN pct_forestation<=75 THEN '50% - 75%'
                ELSE '75% - 100%'
            END AS quartiles
      FROM tab_quartile)

SELECT country_name, region, ROUND(pct_forestation::NUMERIC, 2)
    Pct_Designated_as_Forest
FROM tab_quartile1
WHERE quartiles='75% - 100%'
ORDER BY 1

-- How many countries had a percent forestation higher than the
    United States in 2016?
WITH tab_quartile AS (
  SELECT country_name,
         region,
         pct_forestation
  FROM forestation
  WHERE year=2016 AND pct_forestation IS NOT NULL
  ORDER BY 2),

  tab_quartile1 AS (
    SELECT country_name,
           region,
           pct_forestation,
           CASE
             WHEN pct_forestation<=25 THEN '0 - 25%'
             WHEN pct_forestation<=50 THEN '25% - 50%'
             WHEN pct_forestation<=75 THEN '50% - 75%'
             ELSE '75% - 100%'
           END AS quartiles
    FROM tab_quartile)

SELECT COUNT(*)
FROM tab_quartile1
WHERE pct_forestation>
(SELECT pct_forestation
FROM tab_quartile1
where country_name='United States');
```

Correct! 94 countries had percent forest area higher than the United States