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 sq km** in 1990. As of 2016, the most recent year for which data was available, that number had fallen to **39,958,245.90 sq km**, a loss of **1,324,449.00 sq km**, or -**3.20%(decrease)**.

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 **1,279,999.9891 sq km**).

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%
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 **527,229.062 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 **79,200 sq km**, 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** 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:

Country	Region	Absolute Forest Area Change
Brazil	Latin America & Caribbean	-541,510.00 sq km
Indonesia	East Asia & Pacific	-282,193.98 sq km
Myanmar	East Asia & Pacific	-107,234.00 sq km
Nigeria	Sub-Saharan Africa	-106,506.00 sq km
Tanzania	Sub-Saharan Africa	-102,320.00 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	-75.45 %
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
0-25%	85
25% - 50%	72 (If world is considered as a country_name, count = 73)
50% - 75%	38
75% - 100%	9

The largest number of countries in 2016 were found in the **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%

5. RECOMMENDATIONS

1. There's an overall decrease in the world forest cover contributed especially by the Sub-Saharan Africa and Latin America & Caribbean region: These countries need focus: Nigeria, Togo, Uganda, Mauritania, Honduras since they had the most decrease in forest area, there's a need to understand what could be causing the decrease over the years.
2. Nigeria is seen to have a decrease in forest area both by sq km and percentage hence need to focus to stop the decline. Recommendation to understand what could be driving the decline.
3. From the world bank data we can see that percentage change is hugely affected by the land area of the country: Countries with low land area such as Iceland have a higher percentage change in forest area.
4. Interesting to note is despite Latin America & Caribbean region having the largest forestation, the forest area decreased in both years.
5. Countries such as China and the United States have an increase in forest area which could be a potential case study for countries in the sub-saharan region that are decreasing.
6. The world bank data had in total 5,889 rows however the data had missing information for instance countries such as Hong Kong SAR, China, Kosovo, Monaco, Qatar, San Marino, Sint Maarten (Dutch part), Belgium, Nauru, South Sudan, Sudan didn't have either forest_area or land_area captured making it difficult to know what percentage forest cover they had. For countries like South Sudan and Sudan, they didn't have any land area data collected for any of the years from 2016 - 2010. We can't be 100% certain especially when pointing out top or bottom countries that had an increase or decrease in forest cover given that we had some missing information.
    - I recommend cross checking some of the data against another data source that could have values captured for some of the missing rows in the World bank data.

APPENDIX: SQL queries used

Appendix 1: Global Situation

Forestation view sql:

```
create view forestation
as
select fa.year,
      r.country_code,
      r.country_name,
      r.region,
      fa.forest_area_sqkm,
      la.total_area_sq_mi*2.59 as total_area_sqkm,
      --calculating percentage of land area designated as forest
      (fa.forest_area_sqkm/(la.total_area_sq_mi*2.59))*100 as
percent_forest_area,
      r.income_group
from forest_area as fa
join land_area as la
  on fa.country_code = la.country_code
  and fa.year = la.year
join regions as r
  on la.country_code = r.country_code
;
```

1.a)

```
--use round() to get 2 decimal places
--use cast() to be rounded to numeric in order to use round()
select
      round(cast(forest_area_sqkm as numeric),2) as total_forest_1990
from forestation
where region = 'World'
 and year = 1990
;
```

1.b)

```
select
      round(cast(forest_area_sqkm as numeric),2) as total_forest_2016
from forestation
where region = 'World'
 and year = 2016
;
```

1.c)

```
--1.select forest area & any other relevant column eg region for the year
1990
--2.select forest area & any other relevant column for the year 2016
--3. Get the difference between the 2 years: That is the change
with forest_area_1990 as (
--calculating forest_area in 1990
        select
               round(cast(forest_area_sqkm as numeric),2) forest_area_90,
               region
        from forestation
        where region = 'World'
        and year = 1990
),
--calculating forest_area in 2016
forest_area_2016 as
(
        select
               round(cast(forest_area_sqkm as numeric),2) forest_area_16,
               region
        from forestation
        where region = 'World'
          and year = 2016
)
--Finding change in forest_area
select
       forest_area_90,
       forest_area_16,
       forest_area_16 - forest_area_90 as difference_forest_area
from forest_area_2016 fa16
join forest_area_1990 fa90
  on fa90.region = fa16.region
;
```

1.d)

```
--follow similar steps in c)1&2 above
--calculate percentage change between the 2 years

--calculating forest_area in 1990
with forest_area_1990 as (
```

```
        select
               round(cast(forest_area_sqkm as numeric),2) forest_area_90,
               region
        from forestation
        where region = 'World'
          and year = 1990
),
--calculating forest_area in 2016
forest_area_2016 as (
       select
              round(cast(forest_area_sqkm as numeric),2) forest_area_16,
              region
       from forestation
       where region = 'World'
         and year = 2016
)
--Finding percentage change
select
       forest_area_90,
       forest_area_16,
       round(((forest_area_16 - forest_area_90)/forest_area_90)*100,2) as
percent_change_forest_area
from forest_area_2016 as fa16
join forest_area_1990 as fa90
   on fa90.region = fa16.region
;
```

1.e)

```
select f.country_name,
      f.total_area_sqkm as total_area_sqkm,
      ABS((f.total_area_sqkm)- (select
                                      fa90.forest_area_sqkm -
fa16.forest_area_sqkm as difference
                                from (
                                       select
                                              f.country_name,
                                              f.forest_area_sqkm
                                       from forestation f
                                       where f.country_name = 'World'
                                         and f.year = 1990
                                     ) as fa90
```

```
                                join (
                                        select
                                              f.country_name,
                                              f.forest_area_sqkm
                                          from forestation f
                                        where f.country_name = 'World'
                                          and f.year = 2016) as fa16
                                   on fa90.country_name =
fa16.country_name)) as diff_fa
   from forestation f
   where f.year = 2016
   order by diff_fa
   LIMIT 1
   ;
```

Appendix 2: Regional Outlook

Regional view sql

```
create view regional_forest_area
as
select year,
      region,
      sum(forest_area_sqkm) as sum_forest_area_sqkm,
      sum(total_area_sqkm) as sum_land_area_sqkm,
      (sum(forest_area_sqkm)/sum(total_area_sqkm))*100 as
percent_forest_area
from forestation
where year in (1990, 2016)
group by region, year
order by year
;
```

2.a)

```
select
     round(cast(percent_forest_area AS numeric),2) as percent_forest_region
from regional_forest_area
where year = 2016
  and region = 'World'
;
```

Query for the highest region

```
--use max() to pick highest and limit to give only 1 row, top most
select
      region,
      sum_land_area_sqkm,
      max(round(cast(percent_forest_area AS numeric),2)) as
highest_percent_forest
from regional_forest_area
where year = 2016
group by region, sum_land_area_sqkm
order by highest_percent_forest desc
limit 1
;
```

Query for the lowest region

```
--use min() to pick highest and limit to give only 1 row, top most
select
      region,
      sum_land_area_sqkm,
      min(round(cast(percent_forest_area AS numeric),2)) as
highest_percent_forest
from regional_forest_area
where year = 2016
group by region, sum_land_area_sqkm
order by highest_percent_forest
limit 1
;
```

2.b)

```
--Follow all the steps in 2a above and change the year to 1990.
select
     round(cast(percent_forest_area AS numeric),2) as percent_forest_region
from regional_forest_area
where year = 1990
  and region = 'World'
;
```

Query for highest region

```
select
```

```
      region,
      sum_land_area_sqkm,
      max(round(cast(percent_forest_area AS numeric),2)) as
highest_percent_forest
from regional_forest_area
where year = 1990
group by region, sum_land_area_sqkm
order by highest_percent_forest desc
limit 1
;
```

Query for lowest region

```
select
      region,
      sum_land_area_sqkm,
      min(round(cast(percent_forest_area AS numeric),2)) as
highest_percent_forest
from regional_forest_area
where year = 1990
group by region, sum_land_area_sqkm
order by highest_percent_forest
limit 1
;
```

2.c)

```
--1. select columns for the year 1990 and 2016 separately
--2. select region, percent_forest_area
--3. Calculate difference, add a column for difference
--4. Filter to show percentage forest area in 1990 that were greater than
in 2016
with forest_area_1990 as (
--select all columns for the year 1990
       select
            *
       from regional_forest_area
       where year = 1990
),
forest_area_2016 as (
--select all columns for the year 2016
       select
            *
       from regional_forest_area
```

```
        where year = 2016
)
--pick out region and forest_area for each region
select
        fa90.region,
        round(cast(fa90.percent_forest_area as numeric),2) as forest_1990,
        round(cast(fa16.percent_forest_area as numeric),2) as forest_2016,
--introduce a difference column to see which rows give me a negative
difference hence decrease
        fa16.percent_forest_area - fa90.percent_forest_area as difference
from forest_area_1990 as fa90
join forest_area_2016 as fa16
  on fa90.region = fa16.region
--use where clause to filter out all the rows where forest area in 90 is
larger
where fa90.percent_forest_area > fa16.percent_forest_area
;
```

Appendix 3: Country_Level Detail

3.a)

```
--1. select all columns for years 1990 & 2016
--2. Filter to both 1990 & 2016 separately
--3. Exclude all the nulls in forest_area column
--4. Exclude country 'World'
--5. Calculate the difference (1990 from 2016)
with forest_area_1990 as (
        select
             *
        from forestation f90
        where year = 1990
--Exclude forest_area_sqkms that are null
          and forest_area_sqkm is not null
--Exclude world, doesn't qualify as a country_name
          and country_name != 'World'
),
forest_area_2016 as (
        select
             *
        from forestation f16
        where year = 2016
          and forest_area_sqkm is not null
```

```
        and country_name != 'World'
)
select
      fa90.country_name,
      fa90.region,
      fa90.forest_area_sqkm as forest_area90,
      fa16.forest_area_sqkm as forest_area16,
--get difference in forest area between 2 years
      round(cast(fa16.forest_area_sqkm - fa90.forest_area_sqkm as
numeric),2) as difference
from forest_area_1990 as fa90
join forest_area_2016 as fa16
  on fa90.country_name = fa16.country_name
order by difference
limit 5
;
```

3.b)

```
--Follow same steps in 3b however calculate percent decrease
((2016-1990/1990)*100)
with forest_area_1990 as (
       select
            *
       from forestation f90
       where year = 1990
--Exclude forest_area_sqkms that are null
        and forest_area_sqkm is not null
--Exclude world, doesn't qualify as a country_name
        and country_name != 'World'
),
forest_area_2016 as (
      select
           *
      from forestation f16
      where year = 2016
        and forest_area_sqkm is not null
        and country_name != 'World'
)
select
       fa90.country_name,
       fa90.region,
```

```
       fa90.forest_area_sqkm as forest_area90,
       fa16.forest_area_sqkm as forest_area16,
--get difference in percent forest area between 2 years
       fa16.forest_area_sqkm - fa90.forest_area_sqkm as difference,
       round(cast(((fa16.forest_area_sqkm -
fa90.forest_area_sqkm)/fa90.forest_area_sqkm)*100 as numeric),2) as
perc_difference
from forest_area_1990 as fa90
join forest_area_2016 as fa16
  on fa90.country_name = fa16.country_name
order by perc_difference
limit 5
;
```

3.c)

```
--1. select all columns from forestation db
--2. quartiles are into 4 hence group according to percentage. use case
when to differentiate the quarters
--3. count(quartile groups)
with forest_area_2016 as
(
       select
            *
       from forestation
       where year = 2016
         and percent_forest_area is not null
         and country_name != 'World'
),
percentile as (
      select
           fa16.country_name,
           case
             when fa16.percent_forest_area <=25 then '0-25%'
             when fa16.percent_forest_area > 25 and
fa16.percent_forest_area <=50 then '25%-50%'
             when fa16.percent_forest_area > 50 and
fa16.percent_forest_area <=75 then '50%-75%'
             else '75%-100%'
          end as quartile
from forest_area_2016 as fa16
)
Select
```

```
      distinct quartile,
      count(quartile) as count_quartile_group
from percentile
group by quartile
order by count_quartile_group desc
;
```

3.d)

```
--1. Follow steps 1 & 2 from 3c above
--2. select country, region and percent_forest_area
with forest_area_2016 as
(
       select
            *
       from forestation
       where year = 2016
         and percent_forest_area is not null
         and country_name != 'World'
),
percentile as (
      select
           fa16.*,
           case
             when fa16.percent_forest_area <=25 then '0-25%'
             when fa16.percent_forest_area > 25 and
fa16.percent_forest_area <=50 then '25%-50%'
             when fa16.percent_forest_area > 50 and
fa16.percent_forest_area <=75 then '50%-75%'
            else '75%-100%'
          end as quartile
from forest_area_2016 as fa16
)
select
     distinct p.country_name,
      p.region,
      round(cast(p.percent_forest_area as numeric),2),
      quartile
from percentile p
where quartile = '75%-100%'
;
```

3.e)

```
with forest_area_2016 as (
       select
            *
       from forestation
       where year = 2016
         and percent_forest_area is not null
         and country_name != 'World'
)
select
      count(country_name)
from forest_area_2016 as fa16
where fa16.percent_forest_area > (
       select fa16.percent_forest_area
       from forest_area_2016 fa16
       where fa16.country_name = 'United States'
)
;
```