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 41,282,694.9 sq km in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 39,958,245.9 sq km, a loss of 1,324,449 sq km, or 3.2082%.

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%
Europe & Central Asia	37.28%	38.04%
East Asia & Pacific	25.78%	26.36%
South Asia	16.51%	17.51%
Middle East & North Africa	1.78%	2.07%
North America	35.65%	36.04%

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 527,229.06 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 of 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 5 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 (in sq km)
Brazil	Latin America & Caribbean	541,510
Indonesia	East Asia & Pacific	282,193.98
Myanmar	East Asia & Pacific	107,234
Nigeria	Sub-Saharan Africa	106,506
Tanzania	Sub-Saharan Africa	102,320

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	% 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 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
1	85
2	73
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

4. RECOMMENDATIONS

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

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

  - Amount of forest area decreased from 1990 to 2016, therefore we must raise awareness on this issue so that Earth's natural ecosystem and environment can still retain its natural balance.

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

  - We should focus on *Nigeria* as it is the only country that ranks in the top 5 both in terms of absolute square kilometer decrease in forest and percent decrease in forest area from 1990 to 2016. Therefore, this country has a significant opportunity to stop the decline and any deforestation activities.

  - Moreover, when consider countries that decreased in forest area the most between 1990 and 2016, countries in Sub-Saharan Africa region should be focus on as well as four of the top 5 countries on the list are in this region.

5. APPENDIX: SQL QUERIES USED

Create view called ‘forestation’

```
CREATE VIEW forestation AS (
SELECT      f.country_code,
            f.country_name,
            f.year,
            f.forest_area_sqkm          AS forest_area,
            (l.total_area_sq_mi * 2.59) AS total_land_area,
            r.region,
            r.income_group,
            ROUND(CAST((f.forest_area_sqkm /
            (l.total_area_sq_mi * 2.59)) * 100 AS
            NUMERIC), 2) AS forest_area_percentage
FROM        forest_area     AS f
INNER JOIN  land_area       AS l
ON          f.country_code = l.country_code
AND         f.year = l.year
INNER JOIN  regions         AS r
ON          r.country_code = f.country_code);
```

Part 1: Global situation

a. Find total forest area (in sq km) of the world in 1990 and 2016:

```
SELECT  *
FROM    forestation
WHERE   region = 'World'
AND     year IN (1990, 2016);
```

b. Calculate decreased World's forest area in percentage

```
WITH for_1990 AS
(
SELECT  region,
        forest_area
FROM    forestation
WHERE   region = 'World'
AND     year = 1990), for_2016 AS
(
    SELECT  region,
            forest_area
    FROM    forestation
    WHERE   region = 'World'
    AND     year = 2016)

SELECT          (for_1990.forest_area - for_2016.forest_area)
                AS forest_diff,
                ROUND(CAST(((for_1990.forest_area -
                for_2016.forest_area) * 100/
                for_1990.forest_area) AS NUMERIC), 4)
                AS percent_decrease
FROM            for_1990
INNER JOIN      for_2016
ON              for_1990.region = for_2016.region;
```

c. Comparing the amount of forest area lost between 1990 and 2016, find which country's total area in 2016 is it closest to:

```
SELECT      country_name,
            total_land_area
FROM        forestation
WHERE       year = 2016
AND         total_land_area < 1324449
ORDER BY    total_land_area DESC
LIMIT       1;
```

Part 2: Regional Outlook

Part a) 1. What was the percent forest of the entire world in 2016?

```
WITH forest_1990 AS
(
      SELECT      region,
                  ROUND(CAST((SUM(forest_area) /
                  SUM(total_land_area) *100) AS
                  NUMERIC), 2) AS forest_perc_1990
      FROM        forestation
      WHERE       year = 1990
      GROUP BY    region), forest_2016 AS
(
            SELECT      region,
                        ROUND(CAST((SUM(forest_area) /
                        SUM(total_land_area) *100) AS
                        NUMERIC), 2) AS forest_perc_2016
            FROM        forestation
            WHERE       year = 2016
            GROUP BY    region)
```

```
SELECT          f_2016.region,
                f_2016.forest_perc_2016
FROM            forest_1990 AS f_1990
INNER JOIN      forest_2016 AS f_2016
ON              f_1990.region = f_2016.region
WHERE           f_2016.region = 'World';
```

Part a) 2. Which region had the HIGHEST percent forest in 2016?

```
WITH forest_1990 AS
(
    SELECT      region,
                ROUND(CAST((SUM(forest_area) /
                SUM(total_land_area) *100) AS
                NUMERIC), 2) AS forest_perc_1990
    FROM        forestation
    WHERE       year = 1990
    GROUP BY    region), forest_2016 AS
(
        SELECT      region,
                    ROUND(CAST((SUM(forest_area) /
                    SUM(total_land_area) *100) AS
                    NUMERIC), 2) AS forest_perc_2016
        FROM        forestation
        WHERE       year = 2016
        GROUP BY    region)

SELECT          f_2016.region,
                f_2016.forest_perc_2016
FROM            forest_1990 AS f_1990
```

```
INNER JOIN      forest_2016 AS f_2016
ON              f_1990.region = f_2016.region
ORDER BY        forest_perc_2016 DESC
LIMIT           1;
```

Part a) 3. Which region had the LOWEST percent forest in 2016?

```
WITH forest_1990 AS
(
    SELECT      region,
                ROUND(CAST((SUM(forest_area) /
                SUM(total_land_area) *100) AS
                NUMERIC), 2) AS forest_perc_1990
    FROM        forestation
    WHERE       year = 1990
    GROUP BY    region), forest_2016 AS
(
        SELECT      region,
                    ROUND(CAST((SUM(forest_area) /
                    SUM(total_land_area) *100) AS
                    NUMERIC), 2) AS forest_perc_2016
        FROM        forestation
        WHERE       year = 2016
        GROUP BY    region)
SELECT          f_2016.region,
                f_2016.forest_perc_2016
FROM            forest_1990 AS f_1990
INNER JOIN      forest_2016 AS f_2016
ON              f_1990.region = f_2016.region
ORDER BY        forest_perc_2016 ASC
LIMIT           1;
```

Part b) 1. What was the percent forest of the entire world in 1990?

```
WITH forest_1990 AS
(
	SELECT		region,
			ROUND(CAST((SUM(forest_area) /
			SUM(total_land_area) *100) AS
			NUMERIC), 2) AS forest_perc_1990
	FROM		forestation
	WHERE		year = 1990
	GROUP BY	region), forest_2016 AS
(
		SELECT		region,
				ROUND(CAST((SUM(forest_area) /
				SUM(total_land_area) *100) AS
				NUMERIC), 2) AS forest_perc_2016
		FROM		forestation
		WHERE		year = 2016
		GROUP BY	region)

SELECT		f_2016.region,
		f_1990.forest_perc_1990
FROM		forest_1990 AS f_1990
INNER JOIN	forest_2016 AS f_2016
ON		f_1990.region = f_2016.region
WHERE		f_1990.region = 'World';
```

Part b) 2. Which region had the HIGHEST percent forest in 1990?

```
WITH forest_1990 AS
(
	SELECT		region,
			ROUND(CAST((SUM(forest_area) /
			SUM(total_land_area) *100) AS
			NUMERIC), 2) AS forest_perc_1990
	FROM		forestation
	WHERE		year = 1990
	GROUP BY	region), forest_2016 AS
(
		SELECT		region,
				ROUND(CAST((SUM(forest_area) /
				SUM(total_land_area) *100) AS
				NUMERIC), 2) AS forest_perc_2016
		FROM		forestation
		WHERE		year = 2016
		GROUP BY	region)

SELECT		f_1990.region,
		f_1990.forest_perc_1990
FROM		forest_1990 AS f_1990
INNER JOIN	forest_2016 AS f_2016
ON		f_1990.region = f_2016.region
ORDER BY	forest_perc_1990 DESC
LIMIT		1;
```

Part b) 3. Which region had the LOWEST percent forest in 1990?

```
WITH forest_1990 AS
(
	SELECT		region,
			ROUND(CAST((SUM(forest_area) /
			SUM(total_land_area) *100) AS
			NUMERIC), 2) AS forest_perc_1990
	FROM		forestation
	WHERE		year = 1990
	GROUP BY	region), forest_2016 AS
(
		SELECT		region,
				ROUND(CAST((SUM(forest_area) /
				SUM(total_land_area) *100) AS
				NUMERIC), 2) AS forest_perc_2016
		FROM		forestation
		WHERE		year = 2016
		GROUP BY	region)

SELECT		f_1990.region,
		f_1990.forest_perc_1990
FROM		forest_1990 AS f_1990
INNER JOIN	forest_2016 AS f_2016
ON		f_1990.region = f_2016.region
ORDER BY	forest_perc_1990 ASC
LIMIT		1;
```

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

```
WITH forest_1990 AS
(
	SELECT		region,
			ROUND(CAST((SUM(forest_area) /
			SUM(total_land_area) *100) AS
			NUMERIC), 2) AS forest_perc_1990
	FROM		forestation
	WHERE		year = 1990
	GROUP BY	region), forest_2016 AS
(
		SELECT		region,
				ROUND(CAST((SUM(forest_area) /
				SUM(total_land_area) *100) AS
				NUMERIC), 2) AS forest_perc_2016
		FROM		forestation
		WHERE		year = 2016
		GROUP BY	region)

SELECT		*
FROM		forest_1990 AS f_1990
INNER JOIN	forest_2016 AS f_2016
ON		f_1990.region = f_2016.region
WHERE		f_1990.forest_perc_1990 >
		f_2016.forest_perc_2016;
```

Part 3: Country-Level Detail

A. SUCCESS STORIES

- Query for the first paragraph:

```
WITH for_1990 AS
(
SELECT  country_name,
        region,
        forest_area
FROM    forestation
WHERE   year = 1990
AND     forest_area IS NOT NULL), for_2016 AS
(
    SELECT  country_name,
            region,
            forest_area
    FROM    forestation
    WHERE   year = 2016
    AND     forest_area IS NOT NULL)

SELECT      for_1990.country_name,
            for_1990.region,
            for_2016.forest_area - for_1990.forest_area AS
            forest_diff
FROM        for_1990
INNER JOIN  for_2016
ON          for_1990.country_name = for_2016.country_name
WHERE       for_1990.country_name != 'World'
ORDER BY    forest_diff DESC
LIMIT       5;
```

- Query for top 5 country with highest total land area in 2019 and 2016
  (for the answer China and United States are of course very large countries in total land area) :

```
WITH land_1990 AS
(
SELECT  country_name,
        region,
        total_land_area
FROM    forestation
WHERE   year = 1990
AND     total_land_area IS NOT NULL), land_2016 AS
(
    SELECT  country_name,
            region,
            total_land_area
    FROM    forestation
    WHERE   year = 2016
    AND     total_land_area IS NOT NULL)

SELECT      land_1990.country_name,
            land_1990.total_land_area AS land1990,
            land_2016.total_land_area AS land2016
FROM        land_1990
INNER JOIN  land_2016
ON          land_1990.country_name =
            land_2016.country_name
WHERE       land_1990.country_name != 'World'
```

```
ORDER BY    land_1990.total_land_area DESC,
            land_2016.total_land_area DESC
LIMIT       5;
```

- Query for the answer of the largest *percent* change in forest area from 1990 to 2016:

```
WITH for_1990 AS
(
SELECT  country_name,
        region,
        forest_area
FROM    forestation
WHERE   year = 1990
AND     forest_area IS NOT NULL), for_2016 AS
(
    SELECT  country_name,
            region,
            forest_area
    FROM    forestation
    WHERE   year = 2016
    AND     forest_area IS NOT NULL)

SELECT      for_1990.country_name,
            for_1990.region,
            ((for_2016.forest_area -
            for_1990.forest_area)/for_1990.forest_area) * 100
            AS forest_diff_pct
FROM        for_1990
INNER JOIN  for_2016
```

```
ON          for_1990.country_name = for_2016.country_name
WHERE       for_1990.country_name != 'World'
ORDER BY    forest_diff_pct DESC
LIMIT       1;
```

B. LARGEST CONCERNS

- Query for table 3.1:

```
WITH for_1990 AS
(
SELECT  country_name,
        region,
        forest_area
FROM    forestation
WHERE   year = 1990
AND     forest_area IS NOT NULL), for_2016 AS
(
    SELECT  country_name,
            region,
            forest_area
    FROM    forestation
    WHERE   year = 2016
    AND     forest_area IS NOT NULL)
SELECT      for_1990.country_name,
            for_1990.region,
            for_2016.forest_area - for_1990.forest_area AS
            forest_diff
FROM        for_1990
INNER JOIN  for_2016
ON          for_1990.country_name = for_2016.country_name
WHERE       for_1990.country_name != 'World'
```

```
ORDER BY    forest_diff ASC
LIMIT       5;
```

- Query for table 3.2:

```
WITH for_1990 AS
(
SELECT  country_name,
        region,
        forest_area
FROM    forestation
WHERE   year = 1990
AND     forest_area IS NOT NULL), for_2016 AS
(
    SELECT  country_name,
            region,
            forest_area
    FROM    forestation
    WHERE   year = 2016
    AND     forest_area IS NOT NULL)

SELECT      for_1990.country_name,
            for_1990.region,
            ROUND(CAST((((for_2016.forest_area -
            for_1990.forest_area)/for_1990.forest_area) * 100)
            AS NUMERIC) , 2) AS forest_diff_pct
FROM        for_1990
INNER JOIN  for_2016
ON          for_1990.country_name = for_2016.country_name
WHERE       for_1990.country_name != 'World'
```

```
ORDER BY    forest_diff_pct ASC
LIMIT       5;
```

C. QUARTILES

- Query for table 3.3:

```
WITH forest_pct_quartile AS
(
SELECT      country_name,
            forest_area_percentage,
            CASE WHEN forest_area_percentage <= 25
            THEN 'Q1'
            WHEN (forest_area_percentage > 25 AND
            forest_area_percentage <= 50) THEN 'Q2'
            WHEN (forest_area_percentage > 50 AND
            forest_area_percentage <= 75) THEN 'Q3'
            ELSE 'Q4' END AS quartile
FROM        forestation
WHERE       forest_area_percentage IS NOT NULL
AND         year = 2016
ORDER BY    country_name)
SELECT      DISTINCT(quartile),
            COUNT(quartile) OVER (PARTITION BY quartile)
            AS count_quartile
FROM        forest_pct_quartile
ORDER BY    count_quartile DESC;
```

- Query for table 3.4:

```
WITH forest_pct_quartile AS
(
SELECT      country_name,
            region,
            forest_area_percentage,
            CASE WHEN forest_area_percentage <= 25
            THEN 'Q1'
            WHEN (forest_area_percentage > 25 AND
            forest_area_percentage <= 50) THEN 'Q2'
            WHEN (forest_area_percentage > 50 AND
            forest_area_percentage <= 75) THEN 'Q3'
            ELSE 'Q4' END AS quartile
FROM        forestation
WHERE       forest_area_percentage IS NOT NULL
AND         year = 2016
ORDER BY    country_name)

SELECT      *
FROM        forest_pct_quartile
WHERE       quartile = 'Q4'
ORDER BY    forest_area_percentage DESC;
```

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

```
SELECT      country_name,
            forest_area_percentage
FROM        forestation
WHERE       forest_area_percentage > (
            SELECT      forest_area_percentage
            FROM        forestation
            WHERE       year = 2016
            AND         country_name = 'United States')
AND         year = 2016
ORDER BY    forest_area_percentage DESC;
```