PAOLA CILTLI SEGUNDO METAY
JULY 31ST, 2020

REPORT FOR FORESTQUERY INTO GLOBAL DEFORESTATION, 1990 - 2016

Udacity SQL Nanodegree Program - Part 2 - Introduction to SQL

1. GLOBAL SITUATION

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.

According to the World Bank, the total forest area of the world was **41,282,694.9 km²** in 1990. As of 2016, the most recent year for which data was available, that number had fallen to **39,958,245.9 km²**, a loss of **1,324,449 km²**, 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 **1,279,999.9891**).

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 & The 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 & The Caribbean, with 51.03%, and the region with the lowest relative forestation was Middle East & North Africa, with 1.78% forestation.

Table 3.1: Top 5 Amount Decrease in Forest Area by Country, 1990 & 2016:

Region	1990 Forest Percentage	2016 Forest Percentage
Latin America & Caribbean	51.03	46.16
Europe & Central Asia	37.28	38.04
North America	35.65	36.04
Sub-Saharan Africa	30.67	28.79
East Asia & Pacific	25.78	26.36
South Asia	16.51	17.51
Middle East & North Africa	1.78	2.07

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 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 The United States, but it only saw an increase of 79,200km² , 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 kilometers 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 Change (km²)
Brazil	Latin America & Caribbean	541,510.00
Indonesia	East Asia & Pacific	282,193.98
Myanmar	East Asia & Pacific	107,234.00
Nigeria	Sub-Saharan Africa	106,506.00
Tanzania	Sub-Saharan Africa	102,320.00

B. LARGEST CONCERNS - continuation

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 4 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
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:

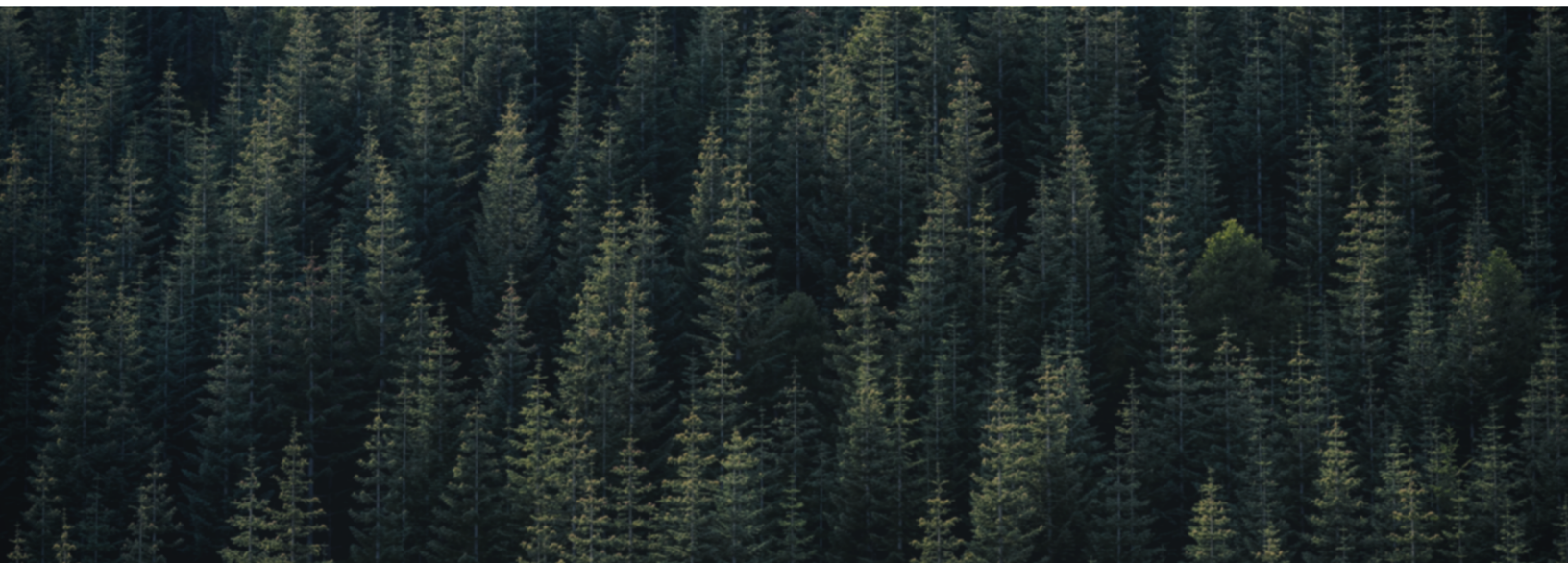
Country	Region	Pct Designated as Forest (%)
Suriname	Latin America & Caribbean	98.26
Micronesia, Fed. Sts.	East Asia & Pacific	91.86
Gabon	Sub-Saharan Africa	90.04
Seychelles	Sub-Saharan Africa	88.41
Palau	East Asia & Pacific	87.61
American Samoa	East Asia & Pacific	87.50
Guyana	Latin America & Caribbean	83.90
Lao PDR	East Asia & Pacific	82.11
Solomon Islands	East Asia & Pacific	77.86

5. RECOMMENDATIONS

The World Bank data on forestation accumulates information by country of the square kilometers of forest per year from 1990 to 2016.
It would be interesting to run the same queries for the up to date data to see if the trend from those years is still current.
As the data shows, the global trend is that the forest coverage is decreasing, which is something to worry and something that must change.

We should focus on the countries that are loosing forest at a fast pace and link that process to other processes that are related to land cover change.

We should also observe closely those countries that are gaining forest area as we could get ideas on how to preserve and manage in a better way forests in other countries.



APPENDIX: SQL queries used

```
# CREATE FORESTATION VIEW
	CREATE VIEW forestation AS
		SELECT forest_area.country_code,
			forest_area.country_name,
			Forest_area.year,
			SUM(forest_area.forest_area_sqkm) AS forest_area,
			AVG((land_area.total_area_sq_mi*2.59)) AS total_area_sqkm,
			SUM(ROUND(CAST((forest_area.forest_area_sqkm*100)/(land_area.total_area
			_sq_mi*2.59)AS NUMERIC),2))AS forest_percent,
			regions.region, regions.income_group
		FROM forest_area
		LEFT JOIN land_area
		ON forest_area.country_code = land_area.country_code
		LEFT JOIN regions
		ON forest_area.country_code = regions.country_code
		GROUP BY forest_area.country_code,
			forest_area.country_name,
			regions.region, regions.income_group, forest_area.year,
			land_area.total_area_sq_mi
		ORDER BY 2 ASC;
```

#1.a What was the total forest area 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_code = 'WLD' AND year = 1990;
```

#1.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 as "World."

```
	SELECT *
	FROM forest_area
	WHERE country_code = 'WLD' AND year = 2016;
```

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

```
	SELECT country_name,
		(SELECT forest_area_sqkm
		FROM forest_area
		WHERE country_code = 'WLD' AND year = 1990) - (
		SELECT forest_area_sqkm
		FROM forest_area
		WHERE country_code = 'WLD' AND year = 2016)AS change_sq_km
	FROM forest_area
	WHERE country_code='WLD'
	GROUP BY country_name;
```

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

```
SELECT country_name,
	100-(((SELECT forest_area_sqkm
		FROM forest_area
		WHERE country_code = 'WLD' AND year = 2016)*100)/
		(SELECT forest_area_sqkm
		FROM forest_area
		WHERE country_code = 'WLD' AND year = 1990))AS change_percentage
FROM forest_area
WHERE country_code='WLD'
GROUP BY country_name;
```

#1.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?

```
SELECT land_area.country_name,
	land_area.year,
	land_area.total_area_sq_mi * 2.59 as total_area_sq_km ,
	ABS(((SELECT forest_area.forest_area_sqkm
		FROM forest_area
		WHERE forest_area.country_code = 'WLD' AND forest_area.year = 1990) -
		(SELECT forest_area.forest_area_sqkm
		FROM forest_area
		WHERE forest_area.country_code = 'WLD' AND forest_area.year = 2016))
		-(land_area.total_area_sq_mi * 2.59)) AS diff
FROM land_area
JOIN forest_area
ON land_area.country_code = forest_area.country_code
WHERE land_area.year = 2016
ORDER BY 4 ASC
LIMIT 1;
```

#2.a.1 What was the percent forest of the entire world in 2016?

```
SELECT land_area.country_code ,
	land_area.year,
	ROUND(CAST((land_area.total_area_sq_mi)*2.59 AS NUMERIC),2) AS
	total_area_sqkm ,
	forest_area.forest_area_sqkm, ROUND(CAST((forest_area.forest_area_sqkm
	*100)/((land_area.total_area_sq_mi)*2.59)AS NUMERIC),2) AS percentage
FROM land_area
JOIN forest_area
ON land_area.country_code = forest_area.country_code
WHERE land_area.country_code = 'WLD'
	AND land_area.year = 2016 AND forest_area.year = 2016;
```

#2.a.2 Which region had the HIGHEST percent forest in 2016

```
SELECT regions.region, land_area.year ,
```

```
        ROUND(CAST(SUM((land_area.total_area_sq_mi)*2.59) AS NUMERIC),2)AS
        total_area_sqkm,
        ROUND(CAST(SUM(forest_area.forest_area_sqkm) AS NUMERIC),2) AS
        forest_sqkm_2016,
        ROUND(CAST((SUM(forest_area.forest_area_sqkm)
        *100)/((SUM(land_area.total_area_sq_mi))*2.59) AS NUMERIC),2) AS percentage
    FROM land_area
    JOIN forest_area
    ON land_area.country_code = forest_area.country_code
    JOIN regions
    ON land_area.country_code = regions.country_code
    WHERE land_area.year = 2016 AND forest_area.year = 2016
    GROUP BY regions.region, land_area.year
    ORDER BY percentage DESC
    LIMIT 1;
```

#2.A.3 which had the LOWEST, to 2 decimal places?

```
    SELECT regions.region, land_area.year ,
        ROUND(CAST(SUM((land_area.total_area_sq_mi)*2.59) AS NUMERIC),2)AS
        total_area_sqkm,
        ROUND(CAST(SUM(forest_area.forest_area_sqkm) AS NUMERIC),2) AS
        forest_sqkm_2016,
        ROUND(CAST((SUM(forest_area.forest_area_sqkm)
        *100)/((SUM(land_area.total_area_sq_mi))*2.59) AS NUMERIC),2) AS percentage
    FROM land_area
    JOIN forest_area
    ON land_area.country_code = forest_area.country_code
    JOIN regions
    ON land_area.country_code = regions.country_code
    WHERE land_area.year = 2016 AND forest_area.year = 2016
    GROUP BY regions.region, land_area.year
    ORDER BY percentage ASC
    LIMIT 1;
```

#2.b.1 What was the percent forest of the entire world in 1990?

```
    SELECT land_area.country_code,
        land_area.year,
        ROUND(CAST((land_area.total_area_sq_mi)*2.59 AS NUMERIC),2) AS
        total_area_sqkm,
        forest_area.forest_area_sqkm,
        ROUND(CAST((forest_area.forest_area_sqkm
        *100)/((land_area.total_area_sq_mi)*2.59)AS NUMERIC),2) AS percentage
    FROM land_area
    JOIN forest_area
```

```
ON land_area.country_code = forest_area.country_code
WHERE land_area.country_code = 'WLD' AND land_area.year = 1990 AND forest_area.year
= 1990;
```

#2.b.2 Which region had the HIGHEST percent forest in 1990

```
SELECT regions.region,
       land_area.year,
       SUM((land_area.total_area_sq_mi)*2.59) AS total_area_sqkm,
       SUM(forest_area.forest_area_sqkm), (SUM(forest_area.forest_area_sqkm)
       *100)/((SUM(land_area.total_area_sq_mi))*2.59) AS percentage
FROM land_area
JOIN forest_area
ON land_area.country_code = forest_area.country_code
JOIN regions
ON land_area.country_code = regions.country_code
WHERE land_area.year = 1990 AND forest_area.year = 1990
GROUP BY regions.region, land_area.year
ORDER BY percentage DESC
LIMIT 1;
```

2.b.3 Which had the LOWEST, to 2 decimal places?

```
SELECT regions.region,
       land_area.year,
       ROUND(CAST(SUM((land_area.total_area_sq_mi)*2.59) AS NUMERIC),2)AS
       total_area_sqkm,
       ROUND(CAST(SUM(forest_area.forest_area_sqkm) AS NUMERIC),2) AS
       forest_sqkm_2016,
       ROUND(CAST((SUM(forest_area.forest_area_sqkm)
       *100)/((SUM(land_area.total_area_sq_mi))*2.59) AS NUMERIC),2) AS percentage
FROM land_area
JOIN forest_area
ON land_area.country_code = forest_area.country_code
JOIN regions
ON land_area.country_code = regions.country_code
WHERE land_area.year = 1990 AND forest_area.year = 1990
GROUP BY regions.region, land_area.year
ORDER BY percentage ASC
LIMIT 1;
```

#2.c.1 Create a table that shows the Regions and their percentage forest area (sum of forest area divided by sum of land area) in 1990 and 2016

```
WITH percentage_1990 AS
       (SELECT regions.region,
```

```
            ROUND(CAST ((SUM(forest_area.forest_area_sqkm)
            *100)/((SUM(land_area.total_area_sq_mi))*2.59) AS NUMERIC),2) AS
            forest_percentage_1990
      FROM land_area
      JOIN forest_area
      ON land_area.country_code = forest_area.country_code
      JOIN regions
      ON land_area.country_code = regions.country_code
      WHERE land_area.year = 1990 AND forest_area.year = 1990
      GROUP BY regions.region, land_area.year),
percentage_2016 AS
      (SELECT regions.region,
            ROUND(CAST ((SUM(forest_area.forest_area_sqkm)
            *100)/((SUM(land_area.total_area_sq_mi))*2.59) AS NUMERIC),2) AS
            Forest_Percentage_2016
      FROM land_area
      JOIN forest_area
      ON land_area.country_code = forest_area.country_code
      JOIN regions
      ON land_area.country_code = regions.country_code
      WHERE land_area.year = 2016 AND forest_area.year = 2016
      GROUP BY regions.region, land_area.year)
SELECT percentage_1990.region , percentage_1990.forest_percentage_1990,
      percentage_2016.forest_percentage_2016
FROM percentage_1990
JOIN percentage_2016
ON percentage_1990.region = percentage_2016.region
WHERE percentage_1990.region <>'World'
ORDER BY 3 DESC;
```

#2.c.2 Based on the table you created, which regions of the world DECREASED in forest area from 1990 to 2016?

```
WITH forest_1990 AS
      (SELECT  regions.region,
            (SUM(forest_area.forest_area_sqkm)
            *100)/((SUM(land_area.total_area_sq_mi))*2.59) AS percentage_1990
      FROM forest_area
      JOIN land_area
      ON forest_area.country_code = land_area.country_code
      JOIN regions
      ON forest_area.country_code = regions.country_code
      WHERE land_area.year = 1990 AND forest_area.year = 1990 AND
            forest_area.country_name<>'World'
```

```
        GROUP BY regions.region),
    forest_2016 AS (
        SELECT regions.region,
            (SUM(forest_area.forest_area_sqkm)*100)/((SUM(land_area.total_area_sq_m
            i))*2.59) AS percentage_2016
        FROM forest_area
        JOIN land_area
        ON forest_area.country_code = land_area.country_code
        JOIN regions
        ON forest_area.country_code = regions.country_code
        WHERE land_area.year = 2016 AND forest_area.year = 2016 AND
            forest_area.country_name<>'World'
        GROUP BY regions.region)
    SELECT regions.region,
        forest_1990.percentage_1990,
        forest_2016.percentage_2016,
        (forest_1990.percentage_1990 -forest_2016.percentage_2016) AS DIFF
    FROM forest_area
    JOIN land_area
    ON forest_area.country_code = land_area.country_code
    JOIN regions
    ON forest_area.country_code = regions.country_code
    JOIN forest_1990
    ON regions.region = forest_1990.region
    JOIN forest_2016
    ON regions.region = forest_2016.region
    WHERE (forest_1990.percentage_1990 -forest_2016.percentage_2016)>0
    GROUP BY regions.region , forest_1990.percentage_1990 , forest_2016.percentage_2016;

#3*SUCCESS STORIES
    WITH forest_1990 AS
        (SELECT country_name ,
            forest_area_sqkm
        FROM  forest_area
        WHERE year = 1990 AND country_name <> 'World' AND forest_area_sqkm
            IS NOT NULL),
    forest_2016 AS
        (SELECT country_name , forest_area_sqkm
        FROM forest_area
        WHERE year = 2016 AND country_name <> 'World' AND forest_area_sqkm
            IS NOT NULL)
    SELECT forest_area.country_name,
        forest_1990.forest_area_sqkm as forest_1990,
```

```
        forest_2016.forest_area_sqkm as forest_2016,
        (forest_2016.forest_area_sqkm-forest_1990.forest_area_sqkm)AS DIFF
    FROM forest_area
    JOIN forest_1990
    ON forest_area.country_name = forest_1990.country_name
    JOIN forest_2016
    ON forest_area.country_name = forest_2016.country_name
    WHERE (forest_2016.forest_area_sqkm-forest_1990.forest_area_sqkm) >0
    GROUP BY forest_area.country_name, forest_1990.forest_area_sqkm ,
    forest_2016.forest_area_sqkm
    ORDER BY DIFF DESC
    LIMIT 5;
```

#3**

```
    WITH forest_1990 AS(
        SELECT country_code , country_name , forest_area , forest_percent , region
        FROM forestation
        WHERE forest_year=1990 AND land_year=1990 AND country_name <> 'World'
            AND forest_area IS NOT NULL AND forest_percent IS NOT NULL
        ORDER BY 4 DESC),
    forest_2016 AS (
        SELECT country_code , country_name , forest_area , forest_percent , region
        FROM forestation
        WHERE forest_year=2016 AND land_year=2016 AND country_name <> 'World'
            AND forest_area IS NOT NULL AND forest_percent IS NOT NULL
        ORDER BY 4 DESC)
    SELECT forest_1990.country_name ,
        forest_1990.region ,
        ROUND(CAST((((forest_2016.forest_area*100)/forest_1990.forest_area)-100)AS
        NUMERIC),2) AS percent_change
    FROM forest_1990
    JOIN forest_2016
    ON forest_1990.country_code=forest_2016.country_code
    WHERE (((forest_2016.forest_area*100)/forest_1990.forest_area)-100)>100
    ORDER BY percent_change DESC
    LIMIT 5;
```

#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 forest_1990 AS(
        SELECT country_code , country_name , forest_area , forest_percent , region
        FROM forestation
```

```
        WHERE forest_year=1990 AND land_year=1990 AND country_name <> 'World' AND
        forest_area IS NOT NULL AND forest_percent IS NOT NULL
        ORDER BY 4 DESC),
    forest_2016 AS (
        SELECT country_code , country_name , forest_area , forest_percent , region
        FROM forestation
        WHERE forest_year=2016 AND land_year=2016 AND country_name <> 'World' AND
        forest_area IS NOT NULL AND forest_percent IS NOT NULL
        ORDER BY 4 DESC)
    SELECT forest_1990.country_name , forest_1990.region ,
        ROUND(CAST((forest_1990.forest_area-forest_2016.forest_area)AS NUMERIC),2)
        AS area_change
    FROM forest_1990
    JOIN forest_2016
    ON forest_1990.country_code=forest_2016.country_code
    WHERE (forest_1990.forest_area-forest_2016.forest_area)>0
    ORDER BY area_change DESC
    LIMIT 5;
```

#3.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 forest_1990 AS(
        SELECT country_code , country_name , forest_area , forest_percent , region
        FROM forestation
            WHERE forest_year=1990 AND land_year=1990 AND country_name <> 'World'
        AND forest_area IS NOT NULL AND forest_percent IS NOT NULL
        ORDER BY 4 DESC),
    forest_2016 AS (
        SELECT country_code , country_name , forest_area , forest_percent , region
        FROM forestation
        WHERE forest_year=2016 AND land_year=2016 AND country_name <> 'World' AND
        forest_area IS NOT NULL AND forest_percent IS NOT NULL
        ORDER BY 4 DESC)
    SELECT forest_1990.country_name ,
        forest_1990.region ,
        ROUND(CAST((100-((forest_2016.forest_area*100)/forest_1990.forest_area))AS
        NUMERIC),2) AS percent_change
    FROM forest_1990
    JOIN forest_2016
    ON forest_1990.country_code=forest_2016.country_code
    WHERE (100-((forest_2016.forest_area*100)/forest_1990.forest_area))>0
    ORDER BY percent_change DESC
    LIMIT 5;
```

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

```
WITH  forest_percent_2016 AS
      (SELECT forest_area.country_name ,
            (SUM(forest_area.forest_area_sqkm)
            *100)/((SUM(land_area.total_area_sq_mi))*2.59)as forest_percent
      FROM forest_area
      JOIN land_area
      ON forest_area.country_code = land_area.country_code
      WHERE land_area.year = 2016 AND forest_area.year = 2016 AND
      forest_area.country_name<>'World' AND forest_area.forest_area_sqkm IS NOT
      NULL AND land_area.total_area_sq_mi IS NOT NULL
      GROUP BY forest_area.country_name) ,
quartiles AS
      (SELECT forest_percent_2016.country_name ,
            CASE WHEN forest_percent_2016.forest_percent <=25 THEN '0-25%'
            WHEN forest_percent_2016.forest_percent<=50 AND
            forest_percent_2016.forest_percent>25 THEN '25%-50%'
            WHEN forest_percent_2016.forest_percent<=75 AND
            forest_percent_2016.forest_percent>50 THEN '50%-75%'
            ELSE '75%-100%'
            END  AS quartiles
      FROM forest_percent_2016)
SELECT quartiles.quartiles , COUNT (quartiles.country_name) AS number_of_countries
FROM quartiles
JOIN forest_percent_2016
ON quartiles.country_name = forest_percent_2016.country_name
GROUP BY quartiles.quartiles
ORDER BY 2 DESC;
```

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

```
SELECT country_name , region , forest_percent
FROM forestation
WHERE forest_year = 2016 AND land_year = 2016 AND forest_percent >75
ORDER BY forest_percent DESC;
```

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

```
WITH  forest_percent_2016 AS
      (SELECT forest_area.country_name ,
            (SUM(forest_area.forest_area_sqkm)
            *100)/((SUM(land_area.total_area_sq_mi))*2.59)as forest_percent
```

```
        FROM forest_area
        JOIN land_area
        ON forest_area.country_code = land_area.country_code
        WHERE land_area.year = 2016 AND forest_area.year = 2016 AND
        forest_area.country_name<>'World' AND forest_area.forest_area_sqkm IS NOT
        NULL AND land_area.total_area_sq_mi IS NOT NULL
        GROUP BY forest_area.country_name)
SELECT COUNT (forest_percent_2016.country_name) AS number_of_countries
FROM forest_percent_2016
WHERE forest_percent_2016.forest_percent >(
        SELECT forest_percent_2016.forest_percent
        FROM forest_percent_2016
        WHERE forest_percent_2016.country_name = 'United States');
```