SQL Queries used in Deforestation Exploration Project

1. GLOBAL SITUATION -

```
CREATE OR REPLACE VIEW forestation AS

SELECT f.country_code,
       f.country_name,
       f.year,
       f.forest_area_sqkm,
       l.total_area_sq_mi,
       r.region,
       r.income_group,
        ((f.forest_area_sqkm*100) / (l.total_area_sq_mi*2.59)) AS forest          /* converting miles
                                                                                     into  sqkm for calculation*/
FROM forest_area f
JOIN land_area l
      ON
          f.country_code = l.country_code
          AND
          f.year = l.year
JOIN regions r
      ON r.country_code = f.country_code
```

••

```
1(A). SELECT forest_area_sqkm
      FROM forestation
      WHERE country_name = 'World'
      And year = '1990'
```

••

```
1(B). SELECT forest_area_sqkm AS total_forest_area
      FROM forestation
      WHERE country_name = 'World'
      And year = '2016'
```

••

```
1(C). WITH loss_table AS
          (SELECT forest_area_sqkm AS total_forest_area,
                 year
           FROM forestation
           WHERE country_name = 'World'
           AND year IN ('1990', '2016'))

      SELECT (t2.total_forest_area - t1.total_forest_area) AS Loss,
             ((t2.total_forest_area - t1.total_forest_area)*100/t2.total_forest_area) AS loss_percent
      FROM loss_table t1
      CROSS JOIN loss_table t2
      WHERE t1.year = '2016'
      AND t2.year = '1990'
```

••

```sql
1(D). CREATE OR REPLACE VIEW loss_amount AS (

    WITH loss_table AS
        (SELECT forest_area_sqkm AS total_forest_area,
                year
         FROM forestation
         WHERE country_name = 'World'
         AND year IN ('1990', '2016'))

    SELECT (t2.total_forest_area - t1.total_forest_area) AS Loss,
           ((t2.total_forest_area - t1.total_forest_area)*100/t2.total_forest_area) AS loss_percent
    FROM loss_table t1
    CROSS JOIN loss_table t2
    WHERE t1.year = '2016'
    AND t2.year = '1990'
    )
```
....
```sql
1(E)   WITH table1 AS (

                        SELECT *
                        FROM land_area
                        WHERE year = '2016'
                        )
    SELECT year,
           country_name,
           (total_area_sq_mi * 2.59) AS Total_area_sqkm,
           ABS((((table1.total_area_sq_mi) * 2.59) - 1324449)) diff_sq_km
    FROM table1
    ORDER BY diff_sq_km
    LIMIT 1
```

2. REGIONAL OUTLOOK
....
```sql
    SELECT region,
           (SUM (forest_area_sqkm)/SUM (total_area_sq_mi*2.59) *100) AS forest_percent,
           year
    FROM forestation
    WHERE year IN ('1990', '2016')
    GROUP BY 1,3
    ORDER BY 1,3 ASC
```

....
3. COUNTRY-LEVEL DETAIL
....
```sql
3(A).   WITH duals1 AS
            (SELECT country_name,
                    forest_area_sqkm
             FROM forest_area
             WHERE year = '2016'
             ORDER BY country_name ASC),
            duals2 AS
            (SELECT country_name,
                    forest_area_sqkm
             FROM forest_area
             WHERE year = '1990'
             ORDER BY country_name ASC)
```

```
SELECT duals1.country_name,
       (duals2.forest_area_sqkm - duals1.forest_area_sqkm) AS difference
FROM duals1
JOIN duals2
ON
duals2.country_name = duals1.country_name
ORDER BY 2 ASC
```

....

```
WITH duals1 AS
        (SELECT country_name,
                forest
         FROM forestation
         WHERE year = '2016'
         ORDER BY country_name ASC),
        duals2 AS
        (SELECT country_name,
                forest
         FROM forestation
         WHERE year = '1990'
         ORDER BY country_name ASC)

SELECT duals1.country_name,
       (duals1.forest - duals2.forest) * 100/ (duals2.forest) AS percentage_change
FROM duals1
JOIN duals2
ON
duals2.country_name = duals1.country_name
ORDER BY 2 DESC
```

....

```
3.1 &3.2)  WITH duals1 AS
        (SELECT country_name,
                region,
                forest_area_sqkm
         FROM forestation
         WHERE year = '2016'
         ORDER BY country_name),
         duals2 AS
         (SELECT country_name,
                 region,
                 forest_area_sqkm
          FROM forestation
          WHERE year = '1990'
          ORDER BY country_name)

SELECT duals1.country_name,
       duals1.region,
       (duals1.forest_area_sqkm - duals2.forest_area_sqkm) AS difference,
       (duals1.forest_area_sqkm - duals2.forest_area_sqkm) *100/duals2.forest_area_sqkm AS percent
FROM duals1
JOIN duals2
ON
duals2.country_name = duals1.country_name
ORDER BY 4 ASC
```

3(C)- 3.3)

```sql
SELECT DISTINCT (quartiles),
        COUNT (country_name) OVER (PARTITION BY quartiles)
FROM (
        SELECT country_name,
                CASE
                        WHEN forest <= 25 THEN '0 - 25%'
                        WHEN forest <= 50 and forest > 25 THEN '25% - 50%'
                        WHEN forest <= 75 and forest > 50 THEN '50% - 75%'
                        ELSE '75% - 100%'
                END AS quartiles
        FROM forestation
        WHERE
        forest IS NOT NULL
        AND
        year = '2016') sub
        order by 1
```

3(C)- 3.4)

```sql
SELECT *
FROM
        (SELECT country_name,
                region,
                forest,
                CASE
                        WHEN forest <= 25 THEN '0 - 25%'
                        WHEN forest <= 50 THEN '25% - 50%'
                        WHEN forest <=75 THEN '50% - 75%'
                        ELSE '75% - 100%'
                END AS quartiles
        FROM forestation
        WHERE forest IS NOT NULL
        AND year = '2016') sub

WHERE quartiles = '75% - 100%'
ORDER BY 3 DESC
```

3(e).

```sql
SELECT COUNT (*)
FROM forestation
WHERE forest >
                (SELECT forest
                 FROM forestation
                 WHERE country_name = 'United States'
                 AND year = '2016')
AND year = '2016'
```

---

5. RECOMMENDATIONS

Write out a set of recommendations as an analyst on the ForestQuery team.
- *What have you learned from the World Bank data?*
- *Which countries should we focus on over others?*

A) My learnings from this dataset are: We can't edit columns, can't delete columns from VIEW. To rerun the VIEW statement, we need to use REPLACE so that we can run our VIEW statement as many times as required.

First time I used CROSS JOIN with WHERE clause and understood how it works practically with the condition.

And learned how to use comments on the workspace. Now, I am very much familiar in using WITH clause. And in the end, I learned how to work on quartiles as well.

B) We need to focus on these countries over others: Togo, Nigeria, Uganda, Mauritania, Honduras