# 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'

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'

)

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**

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**

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)

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)

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).

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?*

1. 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.

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

|  |  |
| --- | --- |
|  |  |