## 1.SQL queries

Tip : Format all SQL queries using <https://sqlformat.org/> for readability

1. GLOBAL SITUATION

**Create forestation view**

CREATE OR REPLACE VIEW FORESTATION AS

SELECT fa.country\_code,

fa.country\_name,

fa.year,

fa.forest\_area\_sqkm,

la.total\_area\_sq\_mi,

r.region,

r.income\_group,

fa.forest\_area\_sqkm/(la.total\_area\_sq\_mi\*2.59)\*100 AS forest\_percentage

FROM forest\_area fa

JOIN land\_area la ON fa.country\_code = la.country\_code

AND fa.year = la.year

JOIN regions r ON fa.country\_code = r.country\_code ;

* 1. **What was the total forest area (in sq km) of the world in 1990? Please keep in mind that you can use the country record denoted as “World" in the region table.**

SELECT forest\_area\_sqkm

FROM forestation

WHERE country\_name = 'World'

AND YEAR = '1990'

* 1. **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 is denoted as “World.”**

SELECT forest\_area\_sqkm

FROM forestation

WHERE country\_name = 'World'

AND YEAR = '2016’

**Combining the two queries in a and b one -**

SELECT country\_name, year, forest\_area\_sqkm

FROM forestation

WHERE country\_name='World' AND (YEAR='1990' OR YEAR='2016')

ORDER BY year ASC;

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

SELECT (f2016.forest\_area\_sqkm - f1990.forest\_area\_sqkm) AS change\_sq\_km

FROM forestation AS f1990, forestation AS f2016

WHERE f2016.year = '2016' AND f2016.country\_name = 'World'

AND f1990.year = '1990' AND f1990.country\_name = 'World';

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

SELECT ((f1990.forest\_area\_sqkm - f2016.forest\_area\_sqkm)/f1990.forest\_area\_sqkm ) \* 100 AS change\_sq\_km

FROM forestation AS f1990, forestation AS f2016

WHERE f1990.year = '1990' AND f1990.country\_name = 'World'

AND f2016.year = '2016' AND f2016.country\_name = 'World';

* 1. **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 country\_name, total\_area\_sq\_mi\*2.59 AS total\_area\_sq\_km

FROM forestation

WHERE year='2016' AND (total\_area\_sq\_mi\*2.59) < 1324449

ORDER BY 2 DESC

LIMIT 1;

**2. REGIONAL OUTLOOK**

**Create the region\_forest\_area View**

CREATE OR REPLACE VIEW region\_forest\_area AS WITH forest\_area\_2016 AS

(SELECT region,

SUM(fa.forest\_area\_sqkm) AS total\_forest\_area\_sqkm\_2016,

SUM(la.total\_area\_sq\_mi \* 2.59) AS total\_area\_sqkm\_2016,

SUM(fa.forest\_area\_sqkm) \* 100 / SUM(la.total\_area\_sq\_mi \* 2.59) AS percent\_fa\_region\_2016

FROM forest\_area fa

JOIN land\_area la ON fa.country\_code = la.country\_code

AND fa.year = la.year

JOIN regions r ON la.country\_code = r.country\_code

WHERE fa.year = 2016

GROUP BY 1

ORDER BY 1),

forest\_area\_1990 AS

(SELECT region,

SUM(fa.forest\_area\_sqkm) AS total\_forest\_area\_sqkm\_1990,

SUM(la.total\_area\_sq\_mi \* 2.59) AS total\_area\_sqkm\_1990,

SUM(fa.forest\_area\_sqkm) \* 100 / SUM(la.total\_area\_sq\_mi \* 2.59) AS percent\_fa\_region\_1990

FROM forest\_area fa

JOIN land\_area la ON fa.country\_code = la.country\_code

AND fa.year = la.year

JOIN regions r ON la.country\_code = r.country\_code

WHERE fa.year = 1990

GROUP BY 1

ORDER BY 1)

SELECT fa2016.region,

total\_forest\_area\_sqkm\_2016,

total\_area\_sqkm\_2016,

percent\_fa\_region\_2016,

total\_forest\_area\_sqkm\_1990,

total\_area\_sqkm\_1990,

percent\_fa\_region\_1990

FROM forest\_area\_2016 fa2016

JOIN forest\_area\_1990 fa1990 ON fa1990.region = fa2016.region;

1. **What was the percent forest of the entire world in 2016? Which region had the HIGHEST percent forest in 2016, and which had the LOWEST, to 2 decimal places?**

SELECT ROUND(percent\_fa\_region\_2016::numeric, 2)

FROM region\_forest\_area

WHERE region = 'World';

SELECT region,

ROUND(percent\_fa\_region\_2016::numeric, 2) AS percent\_fa\_region

FROM region\_forest\_area

WHERE region != 'World'

ORDER BY 2 DESC

LIMIT 1;

SELECT region,

ROUND(percent\_fa\_region\_2016::numeric, 2) AS percent\_fa\_region

FROM region\_forest\_area

WHERE region != 'World'

ORDER BY 2 ASC

LIMIT 1;

1. **What was the percent forest of the entire world in 1990? Which region had the HIGHEST percent forest in 1990, and which had the LOWEST, to 2 decimal places?**

SELECT ROUND(percent\_fa\_region\_1990::numeric, 2)

FROM region\_forest\_area

WHERE region = 'World';

SELECT region,

ROUND(percent\_fa\_region\_1990::numeric, 2) AS percent\_fa\_region

FROM region\_forest\_area

WHERE region != 'World'

ORDER BY 2 DESC

LIMIT 1;

SELECT region,

ROUND(percent\_fa\_region\_1990::numeric, 2) AS percent\_fa\_region

FROM region\_forest\_area

WHERE region != 'World'

ORDER BY 2 ASC

LIMIT 1;

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

SELECT

region,

ROUND(percent\_fa\_region\_1990::numeric, 2) AS percent\_fa\_region\_1990

ROUND(percent\_fa\_region\_2016::numeric, 2) AS percent\_fa\_region\_2016

FROM region\_forest\_area

ORDER BY 2 DESC ;

**3. COUNTRY-LEVEL DETAIL**

**Create the country\_forest\_area view**

CREATE OR REPLACE VIEW country\_forest\_area AS WITH forest\_area\_2016 AS

(SELECT fa.country\_name,

region,

SUM(fa.forest\_area\_sqkm) AS total\_forest\_area\_sqkm\_2016,

SUM(la.total\_area\_sq\_mi \* 2.59) AS total\_area\_sqkm\_2016,

SUM(fa.forest\_area\_sqkm) \* 100 / SUM(la.total\_area\_sq\_mi \* 2.59) AS percent\_fa\_region\_2016

FROM forest\_area fa

JOIN land\_area la ON fa.country\_code = la.country\_code

AND fa.year = la.year

JOIN regions r ON la.country\_code = r.country\_code

WHERE fa.year = 2016

AND fa.forest\_area\_sqkm IS NOT NULL

GROUP BY 1,

2

ORDER BY 1),

forest\_area\_1990 AS

(SELECT fa.country\_name,

region,

SUM(fa.forest\_area\_sqkm) AS total\_forest\_area\_sqkm\_1990,

SUM(la.total\_area\_sq\_mi \* 2.59) AS total\_area\_sqkm\_1990,

SUM(fa.forest\_area\_sqkm) \* 100 / SUM(la.total\_area\_sq\_mi \* 2.59) AS percent\_fa\_region\_1990

FROM forest\_area fa

JOIN land\_area la ON fa.country\_code = la.country\_code

AND fa.year = la.year

JOIN regions r ON la.country\_code = r.country\_code

WHERE fa.year = 1990

AND fa.forest\_area\_sqkm IS NOT NULL

GROUP BY 1,

2

ORDER BY 1)

SELECT fa2016.country\_name "country",

fa2016.region,

total\_forest\_area\_sqkm\_2016,

total\_area\_sqkm\_2016,

percent\_fa\_region\_2016,

total\_forest\_area\_sqkm\_1990,

total\_area\_sqkm\_1990,

percent\_fa\_region\_1990

FROM forest\_area\_2016 fa2016

JOIN forest\_area\_1990 fa1990 ON fa1990.country\_name = fa2016.country\_name;

**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?**

SELECT

country,

region,

ROUND((total\_forest\_area\_sqkm\_1990 - total\_forest\_area\_sqkm\_2016)::Numeric,2) AS "Abs Forest Change"

FROM country\_forest\_area

WHERE country != 'World'

ORDER BY 3 ASC

LIMIT 5;

SELECT

country,

region,

ROUND((total\_forest\_area\_sqkm\_1990 - total\_forest\_area\_sqkm\_2016)::Numeric,2) AS "Abs Forest Change"

FROM country\_forest\_area

WHERE country != 'World'

ORDER BY 3 DESC

LIMIT 5;

**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?**

SELECT

country,

region,

ROUND(((total\_forest\_area\_sqkm\_1990 - total\_forest\_area\_sqkm\_2016)/total\_forest\_area\_sqkm\_1990 \* 100)::Numeric,2) AS "% Forest Change"

FROM country\_forest\_area

WHERE country != 'World'

ORDER BY 3 DESC

LIMIT 5;

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

WITH C1 AS

(SELECT country,

percent\_fa\_region\_2016

FROM country\_forest\_area

GROUP BY 1,

2)

SELECT Distinct(quartiles),

count(country)Over(PARTITION BY quartiles),

percent\_fa\_region\_2016

FROM

(SELECT country,

CASE

WHEN percent\_fa\_region\_2016<25 THEN '1st quartile , 0-25'

WHEN percent\_fa\_region\_2016>=25

AND percent\_fa\_region\_2016<50 THEN '2nd quartile, 25-50'

WHEN percent\_fa\_region\_2016>=50

AND percent\_fa\_region\_2016<75 THEN '3rd quartile, 50-75'

ELSE '4th quartile, 75-100'

END AS quartiles

FROM C1

WHERE percent\_fa\_region\_2016 IS NOT NULL ) q

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

WITH C2 AS

(WITH C1 AS

(SELECT country,

region,

percent\_fa\_region\_2016

FROM country\_forest\_area

GROUP BY 1,

2,

3) SELECT Distinct(quartiles),

count(country)Over(PARTITION BY quartiles),

country,

region,

percent\_fa\_region\_2016

FROM

(SELECT country,

region,

percent\_fa\_region\_2016,

CASE

WHEN percent\_fa\_region\_2016 <=25 THEN '0-25'

WHEN percent\_fa\_region\_2016 >25

AND percent\_fa\_region\_2016 <=50 THEN '25-50'

WHEN percent\_fa\_region\_2016 >50

AND percent\_fa\_region\_2016 <=75 THEN '50-75'

ELSE '75-100'

END AS quartiles

FROM C1

WHERE percent\_fa\_region\_2016 IS NOT NULL ) sub)

SELECT country,

region,

quartiles,

Round(percent\_fa\_region\_2016::Numeric, 2) percent\_forestation

FROM C2

WHERE quartiles = '75-100'

ORDER BY percent\_fa\_region\_2016 DESC