

To see the preview in Atom, you have to tap ctrl+shift+m
Exercices from Udacity SQL Nanodegree -> SQL aggregations

Deforestation Project

Create a View called "forestation" by joining all three tables - forest_area, land_area and regions in the workspace.

The forest_area and land_area tables join on both country_code AND year. The regions table joins these based on only country_code.

In the 'forestation' View, include the following:

All of the columns of the origin tables A new column that provides the percent of the land area that is designated as forest.

Keep in mind that the column forest_area_sqkm in the forest_area table and the land_area_sqmi in the land_area table are in different units (square kilometers and square miles, respectively), so an adjustment will need to be made in the calculation you write (1 sq mi = 2.59 sq km).

```
DROP VIEW IF EXISTS forestation;
CREATE VIEW forestation AS
SELECT
    f.country_code code,
    f.country_name country,
    f.year "year",
    f.forest_area_sqkm forest_area_sqkm,
    l.total_area_sq_mi total_area_sq_mi,
    r.region region,
    r.income_group income_group,
    100.0*(f.forest_area_sqkm /
    (l.total_area_sq_mi * 2.59)) AS perc_forest_per_land
FROM forest_area f
JOIN land_area l ON (f.country_code = l.country_code
AND f.year = l.year)
INNER JOIN regions r ON (f.country_code = r.country_code);
```

```
SELECT * FROM forestation
```

Global Situation

a. 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 forest_area f
WHERE f.country_name = 'World' AND f.year = 1990
```

41282694.9 sqkm 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 is denoted as "World."

```
SELECT forest_area_sqkm FROM forest_area f
WHERE f.country_name = 'World' AND f.year = 2016
```

39958245.9 km2

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

```

WITH forest_1990 AS
(
    SELECT forest_area_sqkm
    FROM forest_area f
    WHERE f.country_name = 'World' AND f.year = 1990
),forest_2016 AS
(
    SELECT forest_area_sqkm
    FROM forest_area f
    WHERE f.country_name = 'World' AND f.year = 2016
)

SELECT (SELECT COALESCE(forest_area_sqkm, 0) FROM forest_1990) - (SELECT COALESCE(forest_area_sqkm, 0) FROM forest_2016) AS differen

```

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

```

WITH forest_1990 AS
(
    SELECT forest_area_sqkm
    FROM forest_area f
    WHERE f.country_name = 'World' AND f.year = 1990
),forest_2016 AS
(
    SELECT forest_area_sqkm
    FROM forest_area f
    WHERE f.country_name = 'World' AND f.year = 2016
)

SELECT 100*((SELECT COALESCE(forest_area_sqkm, 0) FROM forest_1990) - (SELECT COALESCE(forest_area_sqkm, 0) FROM forest_2016))/(SELE

```

3.2 % 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 country, (total_area_sq_mi * 2.59) AS total_area_sqkm
FROM forestation
WHERE year = 2016
ORDER BY total_area_sqkm DESC;

```

Peru 1279999.9891

Regional Outlook

Instructions:

Answering these questions will help you add information into the template.
Use these questions as guides to write SQL queries.

Use the output from the query to answer these questions.

Create a table that shows the Regions and their percent forest area (sum of forest area divided by sum of land area) in 1990 and 2016.
Based on the table you created,

```

...
SELECT
    region,
    perc_for_land,
    f.*
FROM forestation f
WHERE year = 2016 OR year = 1990
...

```

a. 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
    region,
    ROUND(perc_for_land::numeric, 2) percentage
FROM forestation f
WHERE year = 2016 and country = 'World'

```

```

SELECT
    region,
    ROUND(perc_for_land::numeric, 2) percentage
FROM forestation f
WHERE year = 2016
ORDER BY 2 DESC

```

31.38%
 Latin America & Caribbean 98.26
 Europe & Central Asia 0.00

b. 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
    region,
    ROUND(perc_for_land::numeric, 2) percentage
FROM forestation f
WHERE year = 1990
ORDER BY 2 DESC

```

World 32.42%
 Latin America & Caribbean 98.91%
 Europe & Central Asia 0.00%

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

```

WITH region_percent AS
(
    SELECT
        f1990.region,
        SUM(f1990.forest_area_sqkm) region_forest_1990,
        SUM(f1990.total_area_sq_mi * 2.59) region_area_1990,
        SUM(f2016.forest_area_sqkm) region_forest_2016,
        SUM(f2016.total_area_sq_mi * 2.59) region_area_2016
    FROM forestation f1990, forestation f2016
    WHERE f1990.year = '1990'
    AND f2016.year = '2016'
    AND f1990.region = f2016.region
    GROUP BY f1990.region
), forest_perc_90_16 AS
(
    SELECT
        region,
        ROUND(CAST((region_forest_1990/ region_area_1990) * 100 AS NUMERIC), 2)
        AS forest_percent_1990,
        ROUND(CAST((region_forest_2016 / region_area_2016) * 100 AS NUMERIC), 2)
        AS forest_percent_2016
    FROM region_percent
    ORDER BY forest_percent_1990 DESC
)

SELECT fp.*,
fp.forest_percent_1990 - fp.forest_percent_2016 diff
FROM forest_perc_90_16 fp
ORDER BY 4 DESC

```

region	forest_percent_1990	forest_percent_2016	diff
Latin America & Caribbean	51.03	46.16	4.87
Sub-Saharan Africa	30.67	28.79	1.88
World	32.42	31.38	1.04
Middle East & North Africa	1.78	2.07	-0.29

Country-Level Detail

Instructions:

Answering these questions will help you add information into the template.
Use these questions as guides to write SQL queries.
Use the output from the query to answer these questions.

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
    f1990.country_name,
    f1990.forest_area_sqkm forest_area_1990,
    f2016.forest_area_sqkm forest_area_2016,
    f1990.forest_area_sqkm - f2016.forest_area_sqkm diff
FROM forest_area f1990
JOIN forest_area f2016
ON f1990.year = 1990 AND f2016.year = 2016
AND f1990.country_name = f2016.country_name
ORDER BY 4
```

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
    f1990.country,
    f1990.region,
    f1990.forest_area_sqkm forest_area_1990,
    f2016.forest_area_sqkm forest_area_2016,
    f1990.forest_area_sqkm - f2016.forest_area_sqkm diff
FROM forestation f1990
JOIN forestation f2016
ON f1990.year = 1990 AND f2016.year = 2016
AND f1990.country = f2016.country
WHERE f1990.forest_area_sqkm IS NOT NULL
AND f2016.forest_area_sqkm IS NOT NULL
ORDER BY 5 DESC

SELECT
    f1990.country,
    f1990.region,
    f1990.forest_area_sqkm forest_area_1990,
    f2016.forest_area_sqkm forest_area_2016,
    100*(f1990.forest_area_sqkm - f2016.forest_area_sqkm)/f1990.forest_area_sqkm AS diff_perc
FROM forestation f1990
JOIN forestation f2016
ON f1990.year = 1990 AND f2016.year = 2016
AND f1990.country = f2016.country
WHERE f1990.forest_area_sqkm IS NOT NULL
AND f2016.forest_area_sqkm IS NOT NULL
ORDER BY 5 DESC
```

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

```

WITH countries_cat AS
(
    SELECT COUNTRY,
           perc_forest_per_land,
           CASE WHEN perc_forest_per_land <= 25 THEN '0%-25%'
                WHEN perc_forest_per_land <= 75 AND perc_forest_per_land > 50 THEN '50%-75%'
                WHEN perc_forest_per_land <= 50 AND perc_forest_per_land > 25 THEN '25%-50%'
                ELSE '75%-100%' END AS quartiles
    FROM forestation
    WHERE perc_forest_per_land IS NOT NULL AND year = 2016
)

SELECT distinct(quartiles),
       COUNT(country)
FROM countries_cat
GROUP BY quartiles
ORDER BY 1

```

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

```

WITH countries_cat AS
(
    SELECT country,
           region,
           perc_forest_per_land,
           CASE WHEN perc_forest_per_land <= 25 THEN '0%-25%'
                WHEN perc_forest_per_land <= 50 AND perc_forest_per_land > 25 THEN '25%-50%'
                WHEN perc_forest_per_land <= 75 AND perc_forest_per_land > 50 THEN '50%-75%'
                ELSE '75%-100%' END AS quartiles
    FROM forestation
    WHERE perc_forest_per_land IS NOT NULL AND year = 2016
)

SELECT country,
       region,
       perc_forest_per_land
FROM countries_cat
WHERE perc_forest_per_land > 75
ORDER BY 3 DESC

```

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

```

WITH countries_cat AS
(
    SELECT country,
           region,
           perc_forest_per_land,
           CASE WHEN perc_forest_per_land <= 25 THEN '0%-25%'
                WHEN perc_forest_per_land <= 50 AND perc_forest_per_land > 25 THEN '25%-50%'
                WHEN perc_forest_per_land <= 75 AND perc_forest_per_land > 50 THEN '50%-75%'
                ELSE '75%-100%' END AS quartiles
    FROM forestation
    WHERE perc_forest_per_land IS NOT NULL AND year = 2016
)

SELECT COUNT(*)
FROM countries_cat
WHERE perc_forest_per_land > (SELECT perc_forest_per_land FROM countries_cat WHERE country = 'United States')

```