

In questo documento PDF ho caricato delle QUERY in SQL estraendo delle informazioni

da tre database : “World_data” : contiene diversi tipi di dati dei paesi del mondo.

“Global data on sustainable energy” : si concentra sui temi dell’energia sostenibile facendo una panoramica in un arco di 10 anni per ogni paese.

“Global_Missing_Migrants_dataset” : tratta del tema dell’immigrazione rapportato ad altri dati statistici.

I tre dataset e il foglio Excel con i risultati delle query si trovano nella cartella.

```

/* Query 1:
Ho estratto i paesi del database rapportati all'ultimo anno disponibile con
l'accesso di elettricità in percentuale di popolazione e indicando la presenza di
elettricità dal nucleare. Ho ordinato i risultati in ordine discendente dei dati
dell'energia nucleare lasciando i valori nulli.
Nell'allegato Excel i risultati della query. */

```

```

SELECT
    Entity,
    MAX("Year") AS last_year,
    ROUND("Access to electricity (% of population)", 3) AS
    "Access_to_electricity_(%_of_population)", --uso di ROUND per arrotondare le
    "Electricity from nuclear (TWh)" -- le percentuali in output
FROM
    global_data_on_sustainable_energy gdose
GROUP BY
    Entity
ORDER BY
    "Electricity from nuclear (TWh)" DESC;

```

```

/* Query 2: Ho calcolato i primi 5 paesi che hanno avuto il maggior incremento di
"crescita del PIL" facendo la differenza tra
* il primo e l'ultimo anno disponibile. Il risultato è nel foglio Excel in allegato.
*/

```

```

SELECT
    Entity,
    COALESCE( --funzione COALESCE per inserire in OUTPUT il primo valore disponibile
se uno dei due è zero
        (SELECT gdp_growth FROM global_data_on_sustainable_energy gd1
            WHERE gd1.Entity = gdose.Entity -- La condizione gd3.Entity = gdose.Entity
assicura che la sottoquery prenda in considerazione solo le righe che corrispondono
alla stessa entità (Entity) della riga attuale della query principale.
            AND gd1."Year" = (SELECT MAX("Year") FROM global_data_on_sustainable_energy
gd2 WHERE gd2.Entity = gdose.Entity)
        ), 0
    ) -
    COALESCE(
        (SELECT gdp_growth FROM global_data_on_sustainable_energy gd3
            WHERE gd3.Entity = gdose.Entity
            AND gd3."Year" = (SELECT MIN("Year") FROM global_data_on_sustainable_energy
gd4 WHERE gd4.Entity = gdose.Entity)
        ), 0
    ) AS gdp_growth_difference
FROM
    global_data_on_sustainable_energy gdose
GROUP BY
    Entity
ORDER BY
    gdp_growth_difference DESC LIMIT 5;

```

```
/* Query 3:  
Ho selezionato i primi 10 paesi per numero di morti a causa di violenza, ho incluso  
sia la regione  
geografica che il paese in sè per coprire i valori sconosciuti. I risultati nel file  
Excel allegato */
```

```
SELECT "Region of Incident",  
       "Country of Origin",  
       "Number of Dead",  
       "Cause of Death"  
FROM Global_Missing_Migrants_Dataset  
WHERE "Cause of Death" LIKE 'Violence'  
ORDER BY CAST("Number of Dead" AS INTEGER) DESC -- funzione di conversione in intero  
-- a causa di problemi nell'OUTPUT  
LIMIT 10;
```

```
/* Alcune colonne di cui volevo estrarre i dati erano in formato VARCHAR e non mi era  
possibile svolgere  
correttamente le query per cui ho dovuto modificarle prima togliendo il simbolo  
percentuale. Dopo ho creato delle nuove colonne con la dicitura "converted_" e  
le preimpostate in formato REAL e infine ho trasferito i dati delle colonne originali  
in quelle nuove */
```

```
UPDATE world_data -- rimozione del simbolo %  
SET "Gross primary education enrollment (%)" = REPLACE("Gross primary education  
enrollment (%)", '%', ''),  
    "Population: Labor force participation (%)" = REPLACE("Population: Labor force  
participation (%)", '%', ''),  
    "Unemployment rate" = REPLACE("Unemployment rate", '%', ''),  
    "Agricultural Land( %)" = REPLACE("Agricultural Land( %)", '%', '');
```

```
ALTER TABLE world_data -- aggiunta delle nuove colonne in formato REAL  
ADD COLUMN converted_gross_primary_education_enrollment REAL;
```

```
ALTER TABLE world_data  
ADD COLUMN converted_population_labor_force_participation REAL;
```

```
ALTER TABLE world_data  
ADD COLUMN converted_unemployment_rate REAL;
```

```
ALTER TABLE world_data  
ADD COLUMN converted_agricultural_land REAL;
```

```
UPDATE world_data -- trasferimento dei dati dalle vecchie alle nuove colonne  
SET converted_gross_primary_education_enrollment = "Gross primary education enrollment  
(%)";
```

```
UPDATE world_data  
SET converted_population_labor_force_participation = "Population: Labor force  
participation (%)";
```

```
UPDATE world_data  
SET converted_unemployment_rate = "Unemployment rate";
```

```
UPDATE world_data  
SET converted_agricultural_land = "Agricultural Land( %)";
```

```

/* Query 4 :
In questa query (sono due ma fanno parte dello stesso tema) ho
selezionato
i paesi che hanno il minimo e il massimo tasso di disoccupazione a
confronto
con l'iscrizione lorda all'istruzione primaria */
SELECT
    Country,
    MAX("converted_unemployment_rate"),
    "converted_gross_primary_education_enrollment"
FROM world_data wd
WHERE converted_unemployment_rate != ''; --per escludere i valori nulli
',
-- il tasso di disoccupazione è
il valore principale della query

/* South Africa , converted_unemployed_rate : 28.18 % ,
converted_gross_primary_education_enrollment : 100,9 % */

SELECT
    Country,
    MIN("converted_unemployment_rate"),
    "converted_gross_primary_education_enrollment"
FROM world_data wd
WHERE converted_unemployment_rate != '';

/* Qatar , converted_unemployed_rate : 0.09 % ,
converted_gross_primary_education_enrollment : 103.8 % */

/*Query 5:
In questa query ho selezionato i primi dieci paesi con la superficie in
Km/2 in relazione al tasso di natalità facendo una join di due colonne
tra world_data e global_data_on_sustainable_energy. I risultati
nell'allegato Excel. */

SELECT DISTINCT
    gdose."Entity" AS Country, --inserisco in OUTPUT il paese
corrispondente
    gdose."Land Area(Km2)",
    world_data."Birth Rate"
FROM global_data_on_sustainable_energy gdose
LEFT JOIN world_data ON gdose."Entity" = world_data."Country"
WHERE gdose."Land Area(Km2)" IS NOT NULL
AND world_data."Birth Rate" IS NOT NULL
ORDER BY "Land Area(Km2)" DESC -- risultati ordinati in modo discendente
LIMIT 10;

```

/*Query 6 :

In questa query ho estratto il risultato massimo e minimo della media della crescita del PIL dei paesi in elenco. Ho calcolato la media perchè i vari valori sono suddivisi in 10 anni per ogni paese in elenco */

```
SELECT
    "Entity",
    AVG("gdp_growth") AS average_gdp
FROM
    global_data_on_sustainable_energy_gdose
GROUP BY
    "Entity"
ORDER BY average_gdp DESC
LIMIT 1; -- PIL massimo Myanmar : 9.5317503003
```

```
SELECT
    "Entity",
    AVG("gdp_growth") AS average_gdp
FROM
    global_data_on_sustainable_energy_gdose
GROUP BY
    "Entity"
ORDER BY average_gdp ASC
LIMIT 1; -- PIL minimo Barbados : -0.2409643352
```

/* Query 7:

In questa query ho selezionato i primi 10 paesi del mondo per tasso di natalità e di fertilità ordinandoli in ordine discendente riguardo il tasso di mortalità materna. Come intuibile la maggior parte dei paesi del terzo mondo sono i più alti nella mortalità materna ma che negli altri due valori.

Il risultato nell'allegato Excel */

```
SELECT "Country",
    "Maternal mortality ratio",
    "Birth Rate",
    "Fertility Rate"
FROM world_data wd
WHERE "Birth Rate" != ''
```

```
    AND "Maternal mortality ratio" != ''
    AND "Fertility Rate" != ''
ORDER BY "Maternal mortality ratio" DESC
LIMIT 10;
```

/*Query 9:

In questa query ho selezionato i primi 20 paesi per tasso di natalità ,
tassi di fertilità
e tasso di mortalità materna ordinati in modo ascendente per tasso di
nascita.

Si noti come siano tutti paesi con un maggior grado di urbanizzazione. I
risultati

nell'allegato Excel. */

```
SELECT "Country",
       "Birth Rate",
       "Fertility Rate",
       "Maternal mortality ratio"
FROM world_data wd
WHERE "Birth Rate" != ''
    AND "Maternal mortality ratio" != ''
    AND "Fertility Rate" != ''
ORDER BY "Birth Rate" ASC
LIMIT 20; */
```