**Script progetto SQL**

/\* Come da istruzioni ho scaricato i due file csv: "world data 2023" e "global data on sustainable energy"

Dopo averli “puliti” ho creato le due tabelle seguenti: \*/

create table if not exists world\_data\_2023

(

Country varchar (50) not null primary key,

Density int not null,

Abbreviation char (2),

Agricultural\_Land dec (8,4),

Land\_Area int,

Armed\_Forces\_size int,

Birth\_Rate dec (4,2),

Calling\_Code smallint,

Capital\_or\_Major\_City varchar (30),

Co2\_Emissions int,

CPI dec (8,4),

CPI\_Change dec (8,4),

Currency\_Code char (3),

Fertility\_Rate dec (4,2),

Forested\_Area dec (8,4),

Gasoline\_Price\_$ dec (14,2),

GDP\_$ dec (17,3),

Gross\_primary\_education\_enrollment dec (8,4),

Gross\_tertiary\_education\_enrollment dec (8,6),

Infant\_mortality dec (4,2),

Largest\_city varchar (30),

Life\_expectancy dec (4,2),

Maternal\_mortality\_ratio smallint,

Minimum\_wage dec (7,2),

Official\_language varchar (30),

Out\_of\_pocket\_health\_expenditure dec (8,4),

Physician\_per\_thousand dec (4,2),

Population int,

Population\_labor\_force\_participation dec (8,4),

Tax\_revenue dec (8,4),

Total\_tax\_rate dec (8,4),

Unemployment\_rate dec (8,5),

Urban\_population int,

Latitude dec (11,8),

Longitude dec (11,8)

)

create table if not exists Global\_data\_on\_sustainable\_energy

(

Entity varchar (50),

Year smallint,

Access\_to\_electricity\_pct dec (10,8),

Access\_to\_clean\_fuels\_for\_cooking dec (8,6),

Renewable\_electricity\_generating\_capacity\_per\_capita dec (6,2),

Financial\_flows\_to\_developing\_countries bigint,

Renewable\_energy\_share\_in\_total\_final\_energy\_consumption dec (6,5),

Electricity\_from\_fossil\_fuel\_TWh dec (8,2),

Electricity\_from\_nuclear\_TWh dec (8,2),

Electricity\_from\_renewable\_TWh dec (8,2),

Low\_carbon\_electricity\_pct dec (11,10),

Primary\_energy\_consumption\_per\_capita\_kWh dec (12,5),

Energy\_intensity\_level\_of\_primary\_energy dec (12,4),

Value\_co2\_emissions dec (16,8),

Renewable\_pct\_equivalent\_on\_primary\_energy dec (11,10),

Gdp\_growth dec (13,9),

Gdp\_per\_capita dec (14,8),

Density smallint,

Land\_Area\_mq int,

Latitude dec (11,8),

Longitude dec (11,8)

)

/\*Query sulla sanità, con aspettativa di vita e mortalità infantile, uso row\_number per ordinare i risultati \* /

select country, life\_expectancy, infant\_mortality,

row\_number () over (order by life\_expectancy desc) life\_expect\_desc,

row\_number () over (order by infant\_mortality asc) infant\_mortality\_asc

from world\_data\_2023

where life\_expectancy is not null and infant\_mortality is not null

order by life\_expectancy desc

/\* stessa query ma ordinando per mortalità infantile (in ordine crescente) \*/

select country, life\_expectancy, infant\_mortality,

row\_number () over (order by life\_expectancy desc) life\_expect\_desc,

row\_number () over (order by infant\_mortality asc) infant\_mortality\_asc

from world\_data\_2023

where life\_expectancy is not null and infant\_mortality is not null

order by infant\_mortality asc

/\* Query sulla situazione educativa. Ho considerato i valori relativi all’educazione primaria e terziaria, ordinando i risultati per educazione primaria. \*/

select country, gross\_primary\_education\_enrollment, gross\_tertiary\_education\_enrollment,

row\_number () over (order by gross\_primary\_education\_enrollment desc) primary\_education\_desc,

row\_number () over (order by gross\_tertiary\_education\_enrollment desc) tertiary\_education\_des

from world\_data\_2023

where gross\_primary\_education\_enrollment is not null

and gross\_tertiary\_education\_enrollment is not null

/\* questa è stessa query ma ordinandoli secondo l’educazione terziaria \*/

select country, gross\_primary\_education\_enrollment, gross\_tertiary\_education\_enrollment,

row\_number () over (order by gross\_primary\_education\_enrollment desc) primary\_education\_desc,

row\_number () over (order by gross\_tertiary\_education\_enrollment desc) tertiary\_education\_des

from world\_data\_2023

where gross\_primary\_education\_enrollment is not null

and gross\_tertiary\_education\_enrollment is not null

order by gross\_tertiary\_education\_enrollment desc

/\* Query che mette insieme situazione educativa e sanitaria. Ho usato dense\_rank al posto di Row\_number. \*/

select country, gross\_primary\_education\_enrollment, gross\_tertiary\_education\_enrollment, life\_expectancy,

dense\_rank () over (order by gross\_tertiary\_education\_enrollment desc) tertiary\_education\_des,

dense\_rank () over (order by gross\_primary\_education\_enrollment desc) primary\_education\_desc,

dense\_rank () over (order by life\_expectancy desc) top\_life\_expect,

dense\_rank () over (order by physician\_per\_thousand desc) top\_physician\_per\_thousand

from world\_data\_2023

where gross\_primary\_education\_enrollment is not null

and gross\_tertiary\_education\_enrollment is not null

and life\_expectancy is not null

and physician\_per\_thousand is not null

order by life\_expectancy desc

/\* Query per creare la view con la media del pil pro capite: \*/

create view gdp\_per\_capita\_avg as select entity, year, gdp\_per\_capita, (select cast(avg(gdp\_per\_capita) as dec(12,4)) from global\_data\_on\_sustainable\_energy) “avg\_gdp\_per\_capita”

/\* Una volta create la view, effettuo la join tra essa e la tabella ‘global data’, ho utilizzato “partition by” per suddividere i risultati per anno e ho messo come ulteriore clausola, solo i paesi il cui pil pro capite è superiore alla media \*/

select gp.entity, gp.year, w.life\_expectancy, w.physician\_per\_thousand, gp.gdp\_per\_capita, gp.avg\_gdp\_per\_capita,

row\_number () over (partition by year order by life\_expectancy desc) life\_expect\_desc,

row\_number () over (partition by year order by gdp\_per\_capita desc) gdp\_per\_capita\_desc,

row\_number () over (partition by year order by physician\_per\_thousand desc) physician\_desc

from world\_data\_2023 as w

inner join gdp\_per\_capita\_avg as gp on w.country = gp.entity

where life\_expectancy is not null

and gdp\_per\_capita is not null

and physician\_per\_thousand is not null

and gdp\_per\_capita >= avg\_gdp\_per\_capita

/\* Query con la funzione count, per contare quanti paesi, suddivisi anno per anno, hanno un pil pro capite sopra la media: \*/

select year, count (\*) as country\_count

from gdp\_per\_capita\_avg

where gdp\_per\_capita > avg\_gdp\_per\_capita

group by year

order by year

/\* Sempre usando la funzione count, posso utilizzare la query precedente come Subquery per contare solo gli anni in cui i paesi con un pil sopra la media sono più di 50 \*/

select year, country\_count

from (select year, count (\*) as country\_count

from gdp\_per\_capita\_avg

where gdp\_per\_capita > avg\_gdp\_per\_capita

group by year

order by year)

where country\_count > 50

order by country\_count desc

/\* Query sulla produzione di energia. Ho usato partition by per suddividere i risultati in anni \* /

select entity, year, electricity\_from\_fossil\_fuel\_twh, electricity\_from\_nuclear\_twh,

electricity\_from\_renewable\_twh,

dense\_rank () over (partition by year order by electricity\_from\_renewable\_twh desc) as top\_renewable,

dense\_rank () over (partition by year order by electricity\_from\_fossil\_fuel\_twh desc) as top\_fossil\_fuel,

dense\_rank () over (partition by year order by electricity\_from\_nuclear\_twh desc) as top\_nuclear

from global\_data\_on\_sustainable\_energy

where electricity\_from\_fossil\_fuel\_twh is not null

and electricity\_from\_nuclear\_twh is not null

and electricity\_from\_renewable\_twh is not null

/\* qui ho messo come anno di riferimento il 2020: \*/

select entity, year, electricity\_from\_fossil\_fuel\_twh, electricity\_from\_nuclear\_twh,

electricity\_from\_renewable\_twh,

dense\_rank () over (partition by year order by electricity\_from\_renewable\_twh desc) as top\_renewable,

dense\_rank () over (partition by year order by electricity\_from\_fossil\_fuel\_twh desc) as top\_fossil\_fuel,

dense\_rank () over (partition by year order by electricity\_from\_nuclear\_twh desc) as top\_nuclear

from global\_data\_on\_sustainable\_energy

where electricity\_from\_fossil\_fuel\_twh is not null

and electricity\_from\_nuclear\_twh is not null

and electricity\_from\_renewable\_twh is not null

and year = 2020

/\* Query con produzione di energia da fonti rinnovabili più emissioni di co2 \*/

select entity, year, electricity\_from\_renewable\_twh, value\_co2\_emissions,

row\_number () over (partition by year order by electricity\_from\_renewable\_twh desc) top\_renewable,

row\_number () over (partition by year order by value\_co2\_emissions ) min\_co2\_emissions

from global\_data\_on\_sustainable\_energy

where electricity\_from\_renewable\_twh is not null

and value\_co2\_emissions is not null

/\* Query simile alla precedente ma con aggiunta della media del pil pro capite. Ho selezionato solo i paesi in cui il pil pro capite è superiore alla media: \*/

with gdp\_per\_capita\_avg (avg\_gdp\_per\_capita) as (select cast (avg(gdp\_per\_capita) as dec(12,4)) from global\_data\_on\_sustainable\_energy)

select entity, year, electricity\_from\_renewable\_twh, value\_co2\_emissions, gdp\_per\_capita, avg\_gdp\_per\_capita,

row\_number () over (partition by year order by electricity\_from\_renewable\_twh desc) top\_renewable,

row\_number () over (partition by year order by value\_co2\_emissions ) min\_co2\_emissions

from global\_data\_on\_sustainable\_energy, gdp\_per\_capita\_avg

where electricity\_from\_renewable\_twh is not null

and value\_co2\_emissions is not null

and gdp\_per\_capita >= avg\_gdp\_per\_capita

/\* stessa query ma mettendo in ordine crescente per emissioni di co2 \*/

with gdp\_per\_capita\_avg (avg\_gdp\_per\_capita) as (select cast (avg(gdp\_per\_capita) as dec(12,4)) from global\_data\_on\_sustainable\_energy)

select entity, year, electricity\_from\_renewable\_twh, value\_co2\_emissions, gdp\_per\_capita, avg\_gdp\_per\_capita,

row\_number () over (partition by year order by value\_co2\_emissions ) min\_co2\_emissions,

row\_number () over (partition by year order by electricity\_from\_renewable\_twh desc) top\_renewable

from global\_data\_on\_sustainable\_energy, gdp\_per\_capita\_avg

where electricity\_from\_renewable\_twh is not null

and value\_co2\_emissions is not null

and gdp\_per\_capita >= avg\_gdp\_per\_capita