Progetto SQL di Andrea Laiena

World and Energy

Link alla cartella del progetto contenente lo script .sql

*scaricando l’intera cartella e lanciando lo script dal suo interno si evita la necessità di modificare path all’interno dello script.

I DATASET:

I dataset utilizzati nella seguente analisi offrono informazioni circa l’andamento di diversi indicatori di sviluppo di un paese, o semplicemente dati quantitativi come la popolazione di un paese o la sua area coperta da foreste.
Una lista di tutti gli attributi dei dataset, presi dalla clausola CREATE TABLE dello script, è mostrata nelle due seguenti slide

```sql
CREATE TABLE IF NOT EXISTS sustainable_energy (
    country VARCHAR(255), -- name of the country
    report_year NUMERIC(4), --the year the report (row) refers to
    access_to_electricity NUMERIC(9,6), -- % of population with access to electricity
    access_to_clean_fuels_for_cooking NUMERIC(9,6), --% of population with access to clean fuels for cooking
    renewable_elctricity_generation REAL, --renewable electricity generating capacity, expressed in watts per capita
    financial_flows_to_developing_countries BIGINT, --International finance received for clean energy, expressed in USD (NB the datatype MONEY wa
    renewable_energy_share NUMERIC(9,6), --% Share of renewable energy in the total energy consumption
    electricity_from_fossil_fuels REAL, -- Fossil fuels electricity (TWh) Terawatt Hour
    electricity_from_nuclear REAL, -- Nuclear electricity (TWh) Terawatt Hour
    electricity_from_renewables REAL, --Renewable electricity (TWh) Terawatt Hour
    low_carbon_electricity NUMERIC(9,6), -- % Low carbon electricity
    primary_energy_consumption NUMERIC(13,6), --Per capita primary energy consumption (kWh/person)
    energy_intensity REAL, --Energy supplied to the economy per unit value of economic output. (FIXME: Documentation states the measure unit is (
    co2_emissions NUMERIC(14,6), --CO2 emissions in kiloton (kt)
    percentage_primary_energy_from_renewables NUMERIC(9,6), -- percentage of equivalent primary energy that is derived from renewable sources
    gdp_growth NUMERIC(12,9), -- Gross Domestic Product yearly growth
    gdp_per_capita NUMERIC(12,6), -- Gross Domestic Product per capita
    pop_density SMALLINT, -- Population density measured in persons per square kilometer (P/Km^2)
    land_area INT, -- Total land area of the country in square kilometers (Km^2)
    latitude NUMERIC(9,6), -- Latitude of the country
    longitude NUMERIC(9,6), -- Longitude of the country
    PRIMARY KEY (country, report_year)
);
```

```sql
CREATE TABLE IF NOT EXISTS world_2023(
    country VARCHAR(255), -- Name of the country
    pop_density SMALLINT, -- Population density measured in persons per square kilometer (P/Km^2)
    abbreviation CHAR(2), -- Abbreviation or code representing the country
    agricultural_land NUMERIC(4,2), -- Percentage of land area used for agricultural purposes
    land_area INT, -- Total land area of the country in square kilometers (Km^2)
    armed_forces_size INT, -- Size of the armed forces in the country
    birth_rate NUMERIC(4,2), -- Number of births per 1,000 population per year
    calling_code NUMERIC(4), -- International calling code for the country
    capital_or_major_city VARCHAR(255), -- Name of the capital or major city
    co2_emissions INT, -- Carbon dioxide emissions in tons
    cpi NUMERIC(6,2), -- Consumer Price Index, a measure of inflation and purchasing power
    cpi_change NUMERIC(5,2), -- Percentage change in the Consumer Price Index compared to the previous year
    currency_code CHAR(3), -- Currency code used in the country
    fertility_rate NUMERIC(3,2), -- Average number of children born to a woman during her lifetime
    forested_area NUMERIC(4,2), -- Percentage of land area covered by forests
    gasoline_price NUMERIC(3,2), -- Price of gasoline per liter in USD
    gdp BIGINT, -- Gross Domestic Product, the total value of goods and services produced in the country
    gross_primary_education_enrollment NUMERIC(5,2), -- Gross enrollment ratio for primary education
    gross_tertiary_education_enrollment NUMERIC(5,2), -- Gross enrollment ratio for tertiary education
    infant_mortality NUMERIC(4,1), -- Number of deaths per 1,000 live births before reaching one year of age
    largest_city VARCHAR(255), -- Name of the country's largest city
    life_expectancy NUMERIC (3,1), -- Average number of years a newborn is expected to live
    maternal_mortality_ratio SMALLINT, -- Number of maternal deaths per 100,000 live births
    minimum_wage NUMERIC(4,2), -- Minimum wage level in local currency (NB,FIXME: The dataset documentation states "in local currency" but the data (in the raw dataset) reports
    official_language VARCHAR(255), -- Official language(s) spoken in the country
    out_of_pocket_health_expenditure NUMERIC(5,2), -- Percentage of total health expenditure paid out-of-pocket by individuals
    physicians_per_thousand NUMERIC(4,2), -- Number of physicians per thousand people
    population BIGINT, -- Total population of the country
    labor_force_participation NUMERIC(5,2), -- Percentage of the population that is part of the labor force
    tax_revenue NUMERIC(4,2), -- Tax revenue as a percentage of GDP
    total_tax_rate NUMERIC(5,2), -- Overall tax burden as a percentage of commercial profits
    unemployment_rate NUMERIC(4,2), -- Percentage of the labor force that is unemployed
    urban_population INT, -- Percentage of the population living in urban areas
    latitude NUMERIC(9,6), -- Latitude coordinate of the country's location
    longitude NUMERIC(9,6) -- Longitude coordinate of the country's location
);
```

LE OPERAZIONI:

Le operazioni eseguite all’interno dello script non hanno un vero e proprio filo conduttore. Piuttosto gli obbiettivi sono stati due:
1) Estrarre il maggior numero di informazioni rilevanti dai dati che abbiamo a disposizione
2) Creare un blueprint di querys da utilizzare in futuro per eseguire operazioni simili o complementari a quelle eseguite nel sopracitato script.

LE FUNZIONI E I COSTRUTTI UTILIZZATI:

All’interno dello script troviamo esempi di utilizzo di:
METACOMANDI di PSQL,
CLAUSOLE CREATE, ALTER TABLE, ADD CONSTRAINT
VIEWS
Chiaramente SELECT, AS, ma anche DISTINCT, COALESCE, JOIN, LIMIT,
WHERE, ORDER BY, GROUP BY, HAVING, CASE-WHEN-THEN
Funzioni SUM() MAX() ROUND() COUNT()
CTE

LA PAROLA AL CODICE

Lascio il codice parlare, il tutto è commentato nel dettaglio ma senza essere prolisso. Speriamo che vi piaccia.