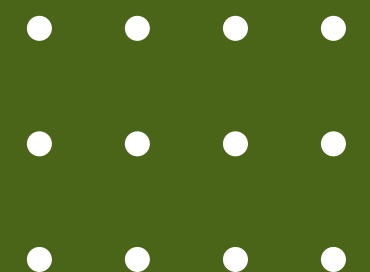


Analisi dei Dati Globali 2023

Un'indagine sugli indicatori economici e ambientali dei
paesi

Progetto SQL Filippo Marzi

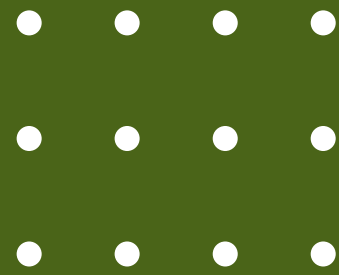


Obiettivo del Progetto:

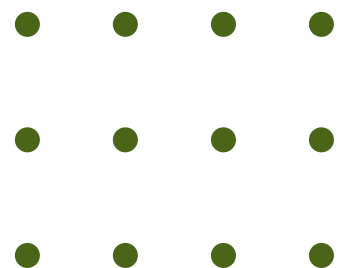
Esplorare le relazioni tra vari indicatori economici e ambientali dei paesi nel 2023, utilizzando un approccio basato sui dati.



Descrizione del Dataset



All'interno del nostro database,
“WorldData2023”, troviamo un'ampia
varietà di informazioni che spaziano
dall'educazione alla sanità, fino ad
arrivare alla fiscalità dei vari paesi.



Premessa



È importante sottolineare che, per questo progetto, ho scelto di concentrarmi esclusivamente sui fattori economici e sulla loro interazione con la sostenibilità ambientale.

Attraverso l'uso del linguaggio SQL, ho voluto non solo mettere in pratica le competenze tecniche acquisite nel corso, ma anche offrire una prospettiva su come i dati possono essere utilizzati per testare le nostre ipotesi

Nel quadro di questo progetto, l'obiettivo non è stato quello di generare grafici per visualizzare i dati, piuttosto quello sviluppare script e utilizzare il linguaggio SQL.

In questo progetto riporterò solamente i risultati più significativi, per maggiori correlazioni o fattori tecnici, lascerò di seguito lo script del codice.

CREAZIONE TABELLA

SCOPO: Conservare e analizzare i dati in modo strutturato per facilitare l'indagine.

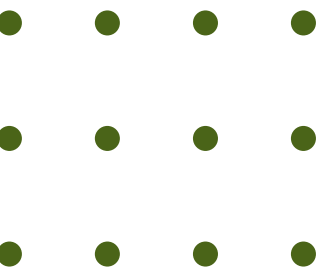
Data Output Messages Notifications							
	country character varying	density integer	abbreviation character varying	agricultural_land numeric	land_area integer	armed_forces_size integer	b n
1	Afghanistan	60	AF	58.10	652230	323000	
2	Albania	105	AL	58.10	28748	9000	
3	Algeria	18	DZ	58.10	2381741	317000	
4	Andorra	164	AD	58.10	468	[null]	
5	Angola	26	AO	58.10	1246700	117000	
6	Antigua and Barbuda	223	AG	58.10	443	0	
7	Argentina	17	AR	58.10	2780400	105000	
8	Armenia	104	AM	58.10	29743	49000	
9	Australia	3	AU	58.10	7741220	58000	
10	Austria	109	AT	58.10	83871	21000	
11	Azerbaijan	123	AZ	58.10	86600	82000	
12	The Bahamas	39	BS	58.10	13880	1000	
13	Bahrain	2239	BH	58.10	765	19000	

IMPORTAZIONE DATI e PRIMA VERIFICA

Iniziamo l'analisi importando i dati nel database WorldData2023. Questo passaggio cruciale assicura che tutte le informazioni siano pronte per l'indagine.

Verifichiamo immediatamente l'integrità dei dati, controllando la presenza di duplicati o valori nulli nelle colonne chiave come country.

Questo mi permette di capire se ho una base dati solida e affidabile.

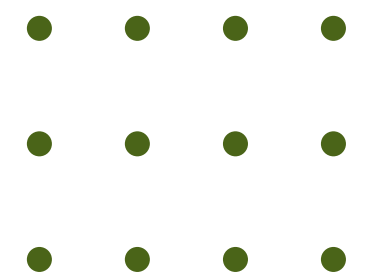


Esplorazione Iniziale dei Dati

La nostra prima indagine rivela che il dataset comprende dati su 195 paesi, senza duplicati o valori nulli nel campo country.

Questa scoperta conferma l'eccellente qualità del nostro dataset, posizionando country come candidato ideale per la chiave primaria

Query	Query History
51	<code>SELECT COUNT (country)</code>
52	<code>FROM WorldData2023;</code>
53	
54	<code>-- La query mi ha dato come output 195, Adesso guardo le città uniche.</code>
55	
56	<code>SELECT COUNT (DISTINCT country)</code>
57	<code>FROM WorldData2023;</code>
58	
59	<code>-- La query mi ha dato come output 195, quindi posso affermare che non ci sono</code>
60	<code>-- Adesso vorrei vedere se ci sono dei valori nulli nelle città.</code>
61	
62	<code>SELECT COUNT (country)</code>
63	<code>FROM WorldData2023</code>
64	<code>WHERE country IS NULL;</code>
65	



Indicatori Economici e Sostenibilità

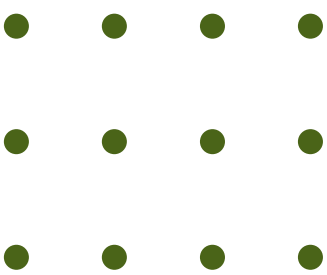
Concentrandoci sugli aspetti economici, esaminiamo il PIL e le emissioni di CO2 per identificare come questi fattori influenzano la sostenibilità aziendale.

Attraverso l'analisi, categorizziamo i paesi in base al loro PIL, confrontandoli con la media globale.

Questo approccio ci permette di evidenziare le disparità economiche e le loro potenziali implicazioni sulla sostenibilità.

Data OutputMessagesNotifications

	country character varying	abbreviation character varying	gdp bigint	gdp_classification text
1	United States	US	21427700000000	gdp ABOVE average
2	China	CN	19910000000000	gdp ABOVE average
3	Japan	JP	5081769542380	gdp ABOVE average
4	Germany	DE	3845630030824	gdp ABOVE average
5	United Kingdom	GB	2827113184696	gdp ABOVE average
6	France	FR	2715518274227	gdp ABOVE average
7	India	IN	26110000000000	gdp ABOVE average
8	South Korea	KR	20290000000000	gdp ABOVE average
9	Italy	IT	2001244392042	gdp ABOVE average
10	Brazil	BR	1839758040766	gdp ABOVE average
11	Canada	CA	1736425629520	gdp ABOVE average
12	Russia	RU	1699876578871	gdp ABOVE average
13	Spain	ES	1394116310769	gdp ABOVE average
14	Australia	AU	1392680589329	gdp ABOVE average



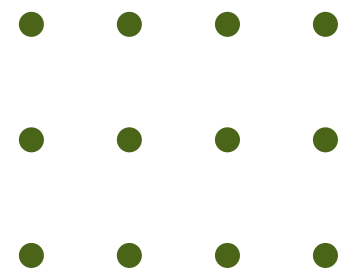
Precisione dei Dati:

Il Caso di São Tomé e Príncipe

Un'attenta revisione rivela un'incongruenza nei dati relativi a São Tomé e Príncipe.

Correggiamo l'errore, assicurando che il nostro dataset sia non solo accurato ma anche rispettoso della realtà geografica e politica di ogni nazione

```
UPDATE WorldData2023
  SET country = 'São Tomé e Príncipe',
      capital_major_city = 'São Tomé',
      largest_city = 'São Tomé'
 WHERE abbreviation = 'ST'
```



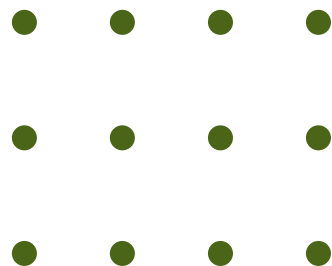
Eliminazione dei Dati:

Vaticano e Palestina

Ho deciso di eliminare Vaticano e Palestina, in quanto entrambe le righe presentavano tutti i valori nulli.

Essendo molto difficile reperire dati attendibili per entrambi, ho preferito direttamente eliminare le righe, così che non inquinassero il database

```
DELETE FROM WorldData2023
WHERE country IN ('Palestinian National Authority', 'Vatican City');
```



Data Output

Messages

Notifications

	country character varying	gdp bigint	tons_co2_emissions integer
1	Iran	445345282123	661710
2	South Africa	351431649241	476644
3	Malaysia	364701517788	248289
4	Kazakhstan	180161741180	247207
5	Egypt	303175127598	238560
6	United Arab Emirates	421142267938	206324
7	Ukraine	153781069118	202250
8	Argentina	449663446954	201348
9	Pakistan	304400000000	201150
10	Vietnam	261921244843	192668
11	Iraq	234094042939	190061

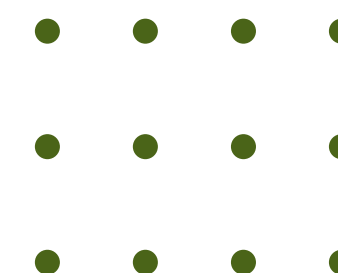
Emissioni di CO2: Un Indicatore Chiave

Approfondiamo l'analisi delle emissioni di CO2, confrontando i dati con il PIL per valutare l'impatto economico sulla sostenibilità.

A prima vista, in termini assoluti, le nazioni con un PIL più alto sono anche quelle che emettono più tonnellate di CO2.

Ma vediamo invece i paesi che con un PIL inferiore alla media, emette CO2 maggiore della media.

Questa tabella rappresenta le nazioni con PIL minori della media ma con emissioni di CO2, in tonnellate, maggiori della media.



Mongolia	13852850259	25368	1.831247687350029
Trinidad and Tobago	24100202834	43868	1.8202336429348245
Turkmenistan	40761142857	70630	1.7327777154773902
Uzbekistan	57921286440	91811	1.585099462442119
Iran	445345282123	661710	1.4858358818702884
Kazakhstan	180161741180	247207	1.3721392698631556
South Africa	351431649241	476644	1.3562921866298205
Ukraine	153781069118	202250	1.3151813884504118
Kyrgyzstan	8454619608	9787	1.157591997484933
Bosnia and Herzegovina	20047848435	21848	1.0897927561072964
Sudan	18902284476	20000	1.0580731670499275
Lesotho	2460072444	2512	1.0211081409926155
Russia	1699876578871	1732027	1.0189133855531756

Emissioni di CO2: Un Indicatore Chiave

Adesso, per avere un paragone più indicativo metterò questi dati non in termini assoluti ma relativi.

Userò come indicatore le tonnellate di CO2 emesse per milione di dollari.

La tabella a fianco rappresenta I nazioni che emettono più CO2 per milione di dollari.











Questo ci indica che i grandi stati, che in termini assoluti emettono più CO2, in realtà, a parità di crescita hanno un’industria e uno sviluppo più sostenibile

PIL Pro Capite e Aspettativa di Vita

Ulteriori analisi dimostrano una correlazione significativa tra il PIL pro capite e l'aspettativa di vita.

Paesi con un PIL superiore alla media tendono a mostrare un'aspettativa di vita maggiore di circa 10 anni, riflettendo come la prosperità economica possa tradursi in benefici tangibili per la salute e il benessere.

PIL pro capite maggiore della media

Data Output		Messages				Notifications			
									
	avg_life_expectancy_above_avg_gdp 								
1	79.7057692307692308								

PIL pro capite minore della media

Data Output

Messages

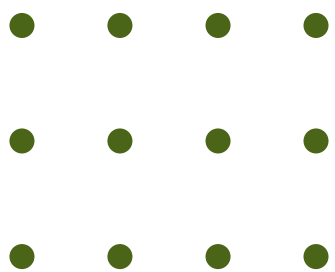
Notifications

avg_life_expectancy_above_avg_gdp

numeric

1

69.4192592592592593



Per approfondire

LINK alla cartella drive con i dettagli tecnici:

https://drive.google.com/drive/folders/1XKOFH7Exuu2p9_a6n0_ebFQAWf-cxdJx?usp=sharing

GRAZIE

Progetto SQL Filippo Marzi

