

# Master of Digital Humanities

## Information Structures and Implications

### Final assignment - Jupyter Notebook report

Students: Geert Hofman & Jacob Moose

---

#### Limitations and Latitudes

*Analyzing possible parameters in global A.I. development*

22 October 2023

## 1. Main research topic

In this project, we investigate whether there are discernable demographic, geographic, socio-economic, political or cultural parameters that impact the successful development of artificial intelligence in a country. Our goal is not to definitively prove the "correct" or "best" route a country should take in order to develop a robust AI sector, but rather to challenge or interrogate criteria currently being used for judging AI and highlight different factors that should possibly be taken into consideration as AI continues to grow.

## 2. Used data sources

In our investigation of the main research question, we collected data from the following sources:

- [Mondial database](#): used for base country data
  - **country**
  - city
  - **economy**
  - **politics**
  - **organization**
  - **ismember**
  - borders
- [Gapminder database](#): used as a shortcut for specific divisions that are generally utilized in country comparisons
  - **ai\_gapminder\_country**
- [Worldbank data](#): used for data on R&D expenditure over the last 20 years
  - **ai\_wb\_data**
- [OECD data](#): used for data surrounding R&D and AI
  - **ai\_oecd\_hiring**
  - **ai\_oecd\_projects\_impact**

- ai\_oecd\_projects\_impact\_pc
- ai\_oecd\_projects\_pop
- ai\_oecd\_projects\_pop\_pc
- ai\_oecd\_pub
- ai\_oecd\_pub\_pc
- **ai\_oecd\_research\_networks**
- ai\_oecd\_scipub\_scopus\_1
- ai\_oecd\_scipub\_scopus\_2
- ai\_oecd\_scipub\_scopus\_3
- ai\_oecd\_talent\_concentration
- [Tortoise media](#): an extensive research project on global AI development that helps establish selective criteria for judging success
  - **ai\_tortoise\_indicators**
- [Our World in Data](#): used for specific data on AI related investments, job postings, and national strategies over time
  - ai\_wid\_investments
  - ai\_wid\_job\_postings
  - ai\_wid\_national\_strategies
- [AI Index](#): contains a lot of data on AI that is used to expand our database
  - aiir\_bills\_2016-2022
  - aiir\_bills\_2022
  - aiir\_hiring\_index\_1
  - aiir\_hiring\_index\_over\_time
  - aiir\_investment\_in\\$\_2013-2022
  - aiir\_investment\_in\\$\_2022
  - aiir\_laws
  - aiir\_leg\_mentions\_2016-2022
  - aiir\_leg\_mentions\_2022
  - aiir\_nat\_strategy\_2017-2022
  - aiir\_new\_aicompanies\_2013-2022
  - aiir\_new\_aicompanies\_2022
  - aiir\_opinions
  - aiir\_skill\_penetration\_rate

The tables in bold were effectively used.

### 3. Methodology

#### ***Organizing and Analyzing the Data***

In order to query this data and produce relevant results, we exported data from our external sources as CSV files and imported them into an adapted MySQL database that was constructed out of the Mondial database. To enhance our performance and have a more workable, integrated system, we added relevant primary and foreign keys to the imported tables. Due to the limitations

of this project and the feedback on the intermediary report, a number of the tables we imported were not utilized during our analysis. Nevertheless, we left them in our database for any future inquiry into our main research topic.

### **Measuring Success**

The central data we used to measure relevant correlations was the Tortoise Media table. This table was constructed based on extensive research by Tortoise Media and serves to establish key criteria for ranking countries' success in AI development. Currently, the table lists 62 different countries with overall "scores" between 0-100 or "ranks" between 1-62. This table serves as a general benchmark for our investigation, as well as something we can critically engage with and against.

### **Analyzing the Data**

The primary link between the data we collect from our different sources is the country code. To ensure that all coding possibilities are readily available, we construct tables that unify data from the Mondial database with information from other tables, such as ai\_gapminder\_country, economy, and politics. It is important to note that this data is not strictly AI-related and contributes to defining a more general profile of each country. From these tables, we create two views that help shape our analysis.

In terms of the questions we ask through our analysis, our attention shifts between three different categories:

1. Economics and Individual Countries
2. Geography, Continents, and World Wealth
3. Global Organizations and Networks.

For our final two queries, several graphics are developed to help reveal certain correlations in more visual ways.

### **Limitations**

While the Tortoise Media table serves as the basis for measuring each country's success in AI development, it is also limited in its subset of countries. In our investigation, relevant countries outside the scope of the Tortoise Media table were also discovered, calling into question the very criteria they propose. This warrants further investigation into common assumptions and considerations made by Tortoise Media. Nevertheless, considering the limited nature of this research project, such realizations were only noted in this report, not deeply investigated.

## **4. Preparing the data**

**Step 1:** Install and import the necessary libraries if needed

Installing libraries

```
In [1]: import sys
```

```
In [ ]: # To install, remove '#'
# Not need to execute these if already installed
# !{sys.executable} -m pip install pandas
# !{sys.executable} -m pip install sqlalchemy
# !{sys.executable} -m pip install sqlalchemy_utils
# !{sys.executable} -m pip install mysql-connector-python
```

Importing libraries and modules

```
In [4]: # For datamanipulation and MySQL connection
import pandas as pd
import sqlalchemy
import sqlalchemy_utils
from sqlalchemy import text, create_engine
from sqlalchemy_utils import database_exists, create_database

# For the graphics
import matplotlib
import numpy as np
import seaborn as sns
```

```
In [5]: import mysql.connector
from mysql.connector import Error

def import_mysqldump(dump_file, host, user, password, database):
    try:
        # Connect to the MySQL database
        connection = mysql.connector.connect(host=host, user=user, password=password, database=database)
        cursor = connection.cursor()

        # Read the SQL file
        with open(dump_file, 'r') as file:
            sql_script = file.read()

        # Execute the SQL script
        for statement in sql_script.split(';'):
            if statement.strip():
                cursor.execute(statement)

        # Commit changes and close the connection
        connection.commit()
        cursor.close()
        connection.close()
        print("Import completed successfully.")

    except Error as e:
        print(f"An error occurred: {e}")
```

**Step 2:** Connect to the database or create it using sqldump

```
In [6]: import getpass
from getpass import getpass

username = input("Enter username:")
passw = getpass("Enter password:")
```

```
In [7]: # hostname = input("Enter host:")
# port = int(input("Enter port to be used:"))
# dbname = input("Enter database name to create or connect to:")

hostname = 'localhost'
port = 3306
```

```
dbname = 'mondial_ai'

connectstring = f'mysql+mysqlconnector://{{username}}:{{passw}}@{{hostname}}:{{port}}/{{dbname}}'
print("You are connecting using the following connectionstring: ", connectstring)
```

You are connecting using the following connectionstring: mysql+mysqlconnector://root:GHafil\_9627@localhost:3306/mondial\_ai

**Step 3:** Check if the database exist; if it does not exist, create it.

```
In [8]: engine = sqlalchemy.create_engine(connectstring)
# If database does not exist, create it and import the sql dump file
if not database_exists(engine.url):
    create_database(engine.url)
    file_path = input("Enter the name of the sql dump file you want to import:")
    import_mysqldump(file_path, hostname, username, passw, dbname)
```

Import completed successfully.

**Step 4:** Use the database

```
In [9]: conn = engine.connect()
print(conn)
query = "USE " + dbname
try:
    conn.execute(text(query))
except Exception as e:
    print(e)

<sqlalchemy.engine.base.Connection object at 0x000001C5AA619350>
```

## 5. Executing the queries

### 5.1. Preparation of base tables

**Query 1:** First, we create a table that unites the country and ai\_gapminder\_country data so that a clear link between country codes is available in this database. To help identify countries worth investigating, we perform a UNION of a right and left join.

```
In [10]: query = """SELECT t1.name, country, code, t2.name, world_4region, g77_and_oecd_countries AS club, income_groups
FROM ai_gapminder_country AS t1
LEFT JOIN country AS t2 ON t1.name = t2.name
UNION
SELECT t1.name, country, code, t2.name, world_4region, g77_and_oecd_countries AS club, income_groups
FROM ai_gapminder_country AS t1
RIGHT JOIN country AS t2 ON t1.name = t2.name
"""

try:
    # Execute the query and store the result in a DataFrame
    df = pd.read_sql_query(query, engine)
    # Display the DataFrame
    print(df.to_markdown()) # This prints the DataFrame in Markdown table format
```

```
except Exception as e:  
    print(e)
```

	name	country	code	name	world_4region	club	income_groups
0	Afghanistan	afg	AFG	Afghanistan	asia	g77	low_income
1	Angola	ago	ANG	Angola	africa	g77	lower_middle_income
2	Albania	alb	AL	Albania	europe	others	upper_middle_income
3	Andorra	and	AND	Andorra	europe	others	high_income
4	United Arab Emirates	are	UAE	United Arab Emirates	asia	g77	high_income
5	Argentina	arg	RA	Argentina	americas	g77	upper_middle_income
6	Armenia	arm	ARM	Armenia	europe	others	upper_middle_income
7	Antigua and Barbuda	atg	AG	Antigua and Barbuda	americas	g77	high_income
8	Australia	aus	AUS	Australia	asia	oecd	high_income
9	Austria	aut	A	Austria	europe	oecd	high_income
10	Azerbaijan	aze	AZ	Azerbaijan	europe	others	upper_middle_income
11	Burundi	bdi	BI	Burundi	africa	g77	low_income
12	Belgium	bel	B	Belgium	europe	oecd	high_income
13	Benin	ben	BEN	Benin	africa	g77	low_income
14	Burkina Faso	bfa	BF	Burkina Faso	africa	g77	low_income
15	Bangladesh	bgd	BD	Bangladesh	asia	g77	lower_middle_income
16	Bulgaria	bgr	BG	Bulgaria	europe	others	upper_middle_income
17	Bahrain	bhr	BRN	Bahrain	asia	g77	high_income
18	Bahamas	bhs	BS	Bahamas	americas	g77	high_income
19	Bosnia and Herzegovina	bih	BIH	Bosnia and Herzegovina	europe	g77	upper_middle_income
20	Belarus	blr	BY	Belarus	europe	others	upper_middle_income
21	Belize	blz	BZ	Belize	americas	g77	upper_middle_income
22	Bolivia	bol	BOL	Bolivia	americas	g77	lower_middle_income
23	Brazil	bra	BR	Brazil	americas	g77	upper_middle_income
24	Barbados	brb	BDS	Barbados	americas	g77	high_income
25	Brunei	brn	BRU	Brunei	asia	g77	high_income
26	Bhutan	btn	BHT	Bhutan	asia	g77	lower_middle_income
27	Botswana	bwa	RB	Botswana	africa	g77	upper_middle_income
28	Central African Republic	caf	RCA	Central African Republic	africa	g77	low_income
29	Canada	can	CDN	Canada	americas	oecd	high_income
30	Switzerland	che	CH	Switzerland	europe	oecd	high_income
31	Chile	chl	RCH	Chile	americas	g77	high_income
32	China	chn	TJ	China	asia	g77	upper_middle_income
33	Cote d'Ivoire	civ			africa	g77	lower_middle_income
34	Cameroon	cmr	CAM	Cameroon	africa	g77	lower_middle_income
35	Congo, Dem. Rep.	cod			africa	g77	low_income
36	Congo, Rep.	cog			africa	g77	lower_middle_income
37	Colombia	col	CO	Colombia	americas	g77	upper_middle_income
38	Comoros	com	COM	Comoros	africa	g77	lower_middle_income
39	Cape Verde	cpv	CV	Cape Verde	africa	g77	lower_middle_income
40	Costa Rica	cri	CR	Costa Rica	americas	g77	upper_middle_income
41	Cuba	cub	C	Cuba	americas	g77	upper_middle_income
42	Cyprus	cyp	CY	Cyprus	europe	others	high_income
43	Czech Republic	cze	CZ	Czech Republic	europe	oecd	high_income
44	Germany	deu	D	Germany	europe	oecd	high_income
45	Djibouti	dji	DJI	Djibouti	africa	g77	lower_middle_income
46	Dominica	dma	WD	Dominica	americas	g77	upper_middle_income
47	Denmark	dnk	DK	Denmark	europe	oecd	high_income
48	Dominican Republic	dom	DOM	Dominican Republic	americas	g77	upper_middle_income
49	Algeria	dza	DZ	Algeria	africa	g77	upper_middle_income
50	Ecuador	ecu	EC	Ecuador	americas	g77	upper_middle_income
51	Egypt	egy	ET	Egypt	africa	g77	lower_middle_income
52	Eritrea	eri	ER	Eritrea	africa	g77	low_income
53	Spain	esp	E	Spain	europe	oecd	high_income

54	Estonia	est	EW	Estonia	europe	others	high_income
55	Ethiopia	eth	ETH	Ethiopia	africa	g77	low_income
56	Finland	fin	SF	Finland	europe	oecd	high_income
57	Fiji	fji	FJI	Fiji	asia	g77	upper_middle_income
58	France	fra	F	France	europe	oecd	high_income
59	Micronesia, Fed. Sts.	fsm			asia	g77	lower_middle_income
60	Gabon	gab	G	Gabon	africa	g77	upper_middle_income
61	United Kingdom	gbr	GB	United Kingdom	europe	oecd	high_income
62	Georgia	geo	GE	Georgia	europe	others	upper_middle_income
63	Ghana	gha	GH	Ghana	africa	g77	lower_middle_income
64	Guinea	gin	RG	Guinea	africa	g77	low_income
65	Gambia	gmb	WAG	Gambia	africa	g77	low_income
66	Guinea-Bissau	gnb	GNB	Guinea-Bissau	africa	g77	low_income
67	Equatorial Guinea	gnq	GQ	Equatorial Guinea	africa	g77	upper_middle_income
68	Greece	grc	GR	Greece	europe	oecd	high_income
69	Grenada	grd	WG	Grenada	americas	g77	upper_middle_income
70	Guatemala	gtm	GCA	Guatemala	americas	g77	upper_middle_income
71	Guyana	guy	GUY	Guyana	americas	g77	upper_middle_income
72	Honduras	hnd	HCA	Honduras	americas	g77	lower_middle_income
73	Holy See	hos	V	Holy See	europe	others	
74	Croatia	hrv	HR	Croatia	europe	others	high_income
75	Haiti	hti	RH	Haiti	americas	g77	low_income
76	Hungary	hun	H	Hungary	europe	oecd	high_income
77	Indonesia	idn	RI	Indonesia	asia	g77	lower_middle_income
78	India	ind	IND	India	asia	g77	lower_middle_income
79	Ireland	irl	IRL	Ireland	europe	oecd	high_income
80	Iran	irn	IR	Iran	asia	g77	upper_middle_income
81	Iraq	irq	IRQ	Iraq	asia	g77	upper_middle_income
82	Iceland	isl	IS	Iceland	europe	oecd	high_income
83	Israel	isr	IL	Israel	asia	oecd	high_income
84	Italy	ita	I	Italy	europe	oecd	high_income
85	Jamaica	jam	JA	Jamaica	americas	g77	upper_middle_income
86	Jordan	jor	JOR	Jordan	asia	g77	upper_middle_income
87	Japan	jpn	J	Japan	asia	oecd	high_income
88	Kazakhstan	kaz			asia	others	upper_middle_income
89	Kenya	ken	EAK	Kenya	africa	g77	lower_middle_income
90	Kyrgyz Republic	kgz			asia	others	lower_middle_income
91	Cambodia	khm	K	Cambodia	asia	g77	lower_middle_income
92	Kiribati	kir	KIR	Kiribati	asia	others	lower_middle_income
93	St. Kitts and Nevis	kna			americas	g77	high_income
94	South Korea	kor	ROK	South Korea	asia	oecd	high_income
95	Kuwait	kwt	KWT	Kuwait	asia	g77	high_income
96	Lao	lao			asia	g77	lower_middle_income
97	Lebanon	lbn	RL	Lebanon	asia	g77	upper_middle_income
98	Liberia	lbr	LB	Liberia	africa	g77	low_income
99	Libya	lby	LAR	Libya	africa	g77	upper_middle_income
100	St. Lucia	lca			americas	g77	upper_middle_income
101	Liechtenstein	lie	FL	Liechtenstein	europe	others	high_income
102	Sri Lanka	lka	CL	Sri Lanka	asia	g77	upper_middle_income
103	Lesotho	lso	LS	Lesotho	africa	g77	lower_middle_income
104	Lithuania	ltu	LT	Lithuania	europe	others	high_income
105	Luxembourg	lux	L	Luxembourg	europe	oecd	high_income
106	Latvia	lva	LV	Latvia	europe	others	high_income
107	Morocco	mar	MA	Morocco	africa	g77	lower_middle_income
108	Monaco	mco	MC	Monaco	europe	others	high_income
109	Moldova	mda	MD	Moldova	europe	others	lower_middle_income

110	Madagascar	mdg	RM	Madagascar	africa	g77	low_income
111	Maldives	mdv	MV	Maldives	asia	g77	upper_middle_income
112	Mexico	mex	MEX	Mexico	americas	oecd	upper_middle_income
113	Marshall Islands	mhl	MH	Marshall Islands	asia	g77	upper_middle_income
114	North Macedonia	mkd			europe	others	upper_middle_income
115	Mali	mli	RMM	Mali	africa	g77	low_income
116	Malta	mlt	M	Malta	europe	others	high_income
117	Myanmar	mmr	MYA	Myanmar	asia	g77	lower_middle_income
118	Montenegro	mne	MNE	Montenegro	europe	others	upper_middle_income
119	Mongolia	mng	MNG	Mongolia	asia	g77	lower_middle_income
120	Mozambique	moz	MOC	Mozambique	africa	g77	low_income
121	Mauritania	mrt	RIM	Mauritania	africa	g77	lower_middle_income
122	Mauritius	mus	MS	Mauritius	africa	g77	upper_middle_income
123	Malawi	mwj	MW	Malawi	africa	g77	low_income
124	Malaysia	mys	MAL	Malaysia	asia	g77	upper_middle_income
125	Namibia	nam	NAM	Namibia	africa	g77	upper_middle_income
126	Niger	ner	RN	Niger	africa	g77	low_income
127	Nigeria	nga	WAN	Nigeria	africa	g77	lower_middle_income
128	Nicaragua	nic	NIC	Nicaragua	americas	g77	lower_middle_income
129	Netherlands	nld	NL	Netherlands	europe	oecd	high_income
130	Norway	nor	N	Norway	europe	oecd	high_income
131	Nepal	npl	NEP	Nepal	asia	g77	low_income
132	Nauru	nru	NAU	Nauru	asia	others	upper_middle_income
133	New Zealand	nzl	NZ	New Zealand	asia	oecd	high_income
134	Oman	omn	OM	Oman	asia	g77	high_income
135	Pakistan	pak	PK	Pakistan	asia	g77	lower_middle_income
136	Panama	pan	PA	Panama	americas	g77	high_income
137	Peru	per	PE	Peru	americas	g77	upper_middle_income
138	Philippines	phl	RP	Philippines	asia	g77	lower_middle_income
139	Palau	plw	PAL	Palau	asia	others	high_income
140	Papua New Guinea	png	PNG	Papua New Guinea	asia	g77	lower_middle_income
141	Poland	pol	PL	Poland	europe	oecd	high_income
142	North Korea	prk	NOK	North Korea	asia	g77	low_income
143	Portugal	prt	P	Portugal	europe	oecd	high_income
144	Paraguay	pry	PY	Paraguay	americas	g77	upper_middle_income
145	Palestine	pse			asia	g77	lower_middle_income
146	Qatar	qat	Q	Qatar	asia	g77	high_income
147	Romania	rou	RO	Romania	europe	others	upper_middle_income
148	Russia	rus	R	Russia	europe	others	upper_middle_income
149	Rwanda	rwa	RWA	Rwanda	africa	g77	low_income
150	Saudi Arabia	sau	SA	Saudi Arabia	asia	g77	high_income
151	Sudan	sdn	SUD	Sudan	africa	g77	lower_middle_income
152	Senegal	sen	SN	Senegal	africa	g77	lower_middle_income
153	Singapore	sgp	SGP	Singapore	asia	g77	high_income
154	Solomon Islands	slb	SLB	Solomon Islands	asia	g77	lower_middle_income
155	Sierra Leone	sle	WAL	Sierra Leone	africa	g77	low_income
156	El Salvador	slv	ES	El Salvador	americas	g77	lower_middle_income
157	San Marino	smr	RSM	San Marino	europe	others	high_income
158	Somalia	som	SP	Somalia	africa	g77	low_income
159	Serbia	srj	SRB	Serbia	europe	others	upper_middle_income
160	South Sudan	ssd			africa		low_income
161	Sao Tome and Principe	stp	STP	Sao Tome and Principe	africa	g77	lower_middle_income
162	Suriname	sur	SME	Suriname	americas	g77	upper_middle_income
163	Slovakia	svk	SK	Slovakia	europe	oecd	high_income
164	Slovenia	svn	SLO	Slovenia	europe	others	high_income
165	Sweden	swe	S	Sweden	europe	oecd	high_income

166	Eswatini	swz	SY	Seychelles	africa	g77	lower_middle_income
167	Seychelles	syc	SYR	Syria	africa	g77	high_income
168	Syria	syr	TCH	Chad	asia	g77	low_income
169	Chad	tcd	RT	Togo	africa	g77	low_income
170	Togo	tgo	THA	Thailand	africa	g77	low_income
171	Thailand	tha	TAD	Tajikistan	asia	g77	upper_middle_income
172	Tajikistan	tjk	TM	Turkmenistan	asia	others	low_income
173	Turkmenistan	tkm	TL	Timor-Leste	asia	g77	upper_middle_income
174	Timor-Leste	tls	TO	Tonga	asia	g77	lower_middle_income
175	Tonga	ton	TT	Trinidad and Tobago	americas	g77	upper_middle_income
176	Trinidad and Tobago	tto	TN	Tunisia	africa	g77	high_income
177	Tunisia	tun	TR	Turkey	europe	oecd	lower_middle_income
178	Turkey	tur	TUV	Tuvalu	asia	others	upper_middle_income
179	Tuvalu	tuv	EAT	Tanzania	africa	g77	upper_middle_income
180	Tanzania	tza	EAU	Uganda	africa	g77	low_income
181	Uganda	uga	UA	Ukraine	europe	g77	low_income
182	Ukraine	ukr	ROU	Uruguay	americas	others	lower_middle_income
183	Uruguay	ury	USA	United States	americas	g77	high_income
184	United States	usa	UZB	Uzbekistan	asia	oecd	high_income
185	Uzbekistan	uzb	vct	Venezuela	americas	others	lower_middle_income
186	St. Vincent and the Grenadines	vct	ven	Venezuela	americas	g77	upper_middle_income
187	Venezuela	ven	VN	Vietnam	asia	g77	upper_middle_income
188	Vietnam	vnm	VU	Vanuatu	asia	g77	lower_middle_income
189	Vanuatu	vut	WS	Samoa	asia	g77	lower_middle_income
190	Samoa	wsm	YE	Yemen	asia	g77	upper_middle_income
191	Yemen	yem	RSA	South Africa	africa	g77	low_income
192	South Africa	zaf	Z	Zambia	africa	g77	upper_middle_income
193	Zambia	zmb	ZW	Zimbabwe	africa	g77	lower_middle_income
194	Zimbabwe	zwe	AMSA	American Samoa		g77	lower_middle_income
195			AXA	Anguilla			
196			ARU	Aruba			
197			BERM	Bermuda			
198			BVIR	British Virgin Islands			
199			CAYM	Cayman Islands			
200			XMAS	Christmas Island			
201			COCO	Cocos Islands			
202			RCB	Congo			
203			COOK	Cook Islands			
204			CI	Cote d'Ivoire			
205			FALK	Falkland Islands			
206			FARX	Faroe Islands			
207			FGU	French Guiana			
208			FPOL	French Polynesia			
209			GAZA	Gaza Strip			
210			GBZ	Gibraltar			
211			GROX	Greenland			
212			GUAD	Guadeloupe			
213			GUAM	Guam			
214			GBG	Guernsey			
215			HONX	Hong Kong			
216			GBJ	Jersey			
217			KAZ	Kazakhstan			
218			KOS	Kosovo			
219			KGZ	Kyrgyzstan			
220			LAO	Laos			

222		MACX	Macau
223		MK	Macedonia
224		GBM	Man
225		MART	Martinique
226		MAYO	Mayotte
227		FSM	Micronesia
228		MNTS	Montserrat
229		NA	Netherlands Antilles
230		NCA	New Caledonia
231		NIUE	Niue
232		NORF	Norfolk Island
233		NMIS	Northern Mariana Islands
234		PITC	Pitcairn Islands
235		PR	Puerto Rico
236		REUN	Reunion
237		HELX	Saint Helena
238		KN	Saint Kitts and Nevis
239		WL	Saint Lucia
240		SMAR	Saint Martin
241		SPMI	Saint Pierre and Miquelon
242		WV	Saint Vincent and the Grenadines
243		SVAX	Svalbard
244		SD	Swaziland
245		RC	Taiwan
246		TUCA	Turks and Caicos Islands
247		VIRG	Virgin Islands
248		WAFU	Wallis and Futuna
249		WEST	West Bank
250		WSA	Western Sahara
251		ZRE	Zaire

**Query 1 Results:** We now have a clear picture of the countries that are included and excluded in the country and ai\_gapminder\_country data. Since we would like to create a view that is based on independent, UN-recognized countries with an identifiable GDP and population, the countries that are located in the country table but not in the ai\_gapminder\_country table (such as Bermuda, Gibraltar, Loas, and more) will most likely be discarded.

**Query 2:** But before we create a view based on independent, UN-recognized countries with an identifiable GDP and population, it is best to verify whether all the countries in the ai\_tortoise\_indicators table are explicitly present in the ai\_gapminder\_table.

```
In [11]: # Check if there are countries in 'ai_tortoise_indicators' that are not in 'ai_gapminder_country'
# by doing a left join on it and checking if there are data available from that second table
query = """SELECT t1.country, t2.country, t2.name, world_4region, g77_and_oecd_countries AS club, income_groups
FROM ai_tortoise_indicators AS t1
LEFT JOIN ai_gapminder_country AS t2 ON t1.country = t2.name
WHERE t2.name IS NULL
"""

try:
    # Execute the query and store the result in a DataFrame
    df = pd.read_sql_query(query, engine)
    # Display the DataFrame
    print(df.to_markdown()) # This prints the DataFrame in Markdown table format
except Exception as e:
    print(e)
```

	country	country	name	world_4region	club	income_groups
0	Hong Kong					
1	Taiwan					

**Query 2 Results:** The query reveals two regions (each officially a part of China, according to the UN) that are included as independent entities in the ai\_tortoise\_indicators table but not in the ai\_gapminder\_country table. This is worth investigating further, and we can check whether these regions have data in other relevant tables. If so, we can add them to our base view.

**Query 3:** By executing a join between the ai\_tortoise\_indicators, country, economy and politics tables, we can verify if there is any relevant data (such as area, population, and GDP) available for these regions.

```
In [12]: query = """SELECT t1.country, t2.name, t2.code, t2.area, t2.population, t3.GDP, t4.dependent, t4.government
FROM ai_tortoise_indicators AS t1
LEFT JOIN country AS t2 ON t1.country = t2.name
LEFT JOIN economy AS t3 ON t2.code = t3.country
LEFT JOIN politics AS t4 ON t2.code = t4.country
"""

try:
    # Execute the query and store the result in a DataFrame
    df = pd.read_sql_query(query, engine)
    # Display the DataFrame
    print(df.to_markdown()) # This prints the DataFrame in Markdown table format
except Exception as e:
    print(e)
```

## ISI-AI-Final-Assignment-report

	country	name	code	area	population	GDP	dependent	government
0	Argentina	Argentina	RA	2.76689e+06	34672997   278500			republic
1	Armenia	Armenia	ARM	29800	3463574   9100			republic
2	Australia	Australia	AUS	7.68685e+06	18260863   405400			federal parliamentary state
3	Austria	Austria	A	83850	8023244   152000			federal republic
4	Bahrain	Bahrain	BRN	620	590042   7300			traditional monarchy
5	Belgium	Belgium	B	30510	10170241   197000			constitutional monarchy
6	Brazil	Brazil	BR	8.51196e+06	162661214   976800			federal republic
7	Canada	Canada	CDN	9.97614e+06	28820671   694000			confederation with parliamentary democracy
8	Chile	Chile	RCH	756950	14333258   113200			republic
9	China	China	TJ	9.59696e+06	1210004956   3.5e+06			Communist state
10	Colombia	Colombia	CO	1.13891e+06	36813161   192500			republic
11	Czech Republic	Czech Republic	CZ	78703	10321120   106200			parliamentary democracy
12	Denmark	Denmark	DK	43070	5249632   112800			constitutional monarchy
13	Egypt	Egypt	ET	1.00145e+06	63575107   171000			republic
14	Estonia	Estonia	EW	45100	1459428   12300			republic
15	Finland	Finland	SF	337030	5105230   92400			republic
16	France	France	F	547030	58317450   1.173e+06			republic
17	Germany	Germany	D	356910	83536115   1.4522e+06			federal republic
18	Greece	Greece	GR	131940	10538594   101700			parliamentary republic
19	Hong Kong	Hong Kong	HONX	1092	7055071   307600	TJ		special administrative area in China
20	Hungary	Hungary	H	93030	10002541   72500			republic
21	Iceland	Iceland	IS	103000	270292   5000			republic
22	India	India	IND	3.28759e+06	952107694   1.4087e+06			federal republic
23	Indonesia	Indonesia	RI	1.91944e+06	206611600   710900			republic
24	Ireland	Ireland	IRL	70280	3566833   54600			republic
25	Israel	Israel	IL	20770	5421995   80100			republic

## ISI-AI-Final-Assignment-report

26   Italy	Italy	I	301230	57460274	1.0886e+06	republic
27   Japan	Japan	J	377835	125449703	2.6792e+06	constitutional monarchy
28   Kenya	Kenya	EAK	582650	28176686	36800	republic
29   Lithuania	Lithuania	LT	65200	3646041	13300	republic
30   Luxembourg	Luxembourg	L	2586	415870	10000	constitutional monarchy
31   Malaysia	Malaysia	MAL	329750	19962893	193600	constitutional monarchy
32   Malta	Malta	M	320	375576	4400	parliamentary democracy
33   Mexico ized government	Mexico	MEX	1.97255e+06	95772462	721400	federal republic operating under a central
34   Morocco	Morocco	MA	446550	29779156	87400	constitutional monarchy
35   Netherlands	Netherlands	NL	37330	15568034	301900	constitutional monarchy
36   New Zealand	New Zealand	NZ	268680	3547983	62300	parliamentary democracy
37   Nigeria	Nigeria	WAN	923770	103912489	135900	military government
38   Norway	Norway	N	324220	4383807	106200	constitutional monarchy
39   Pakistan	Pakistan	PK	803940	129275660	274200	republic
40   Poland	Poland	PL	312683	38642565	226700	democratic state
41   Portugal	Portugal	P	92080	9865114	116200	republic
42   Qatar	Qatar	Q	11000	547761	10700	traditional monarchy
43   Russia	Russia	R	1.70752e+07	148178487	796000	federation
44   Saudi Arabia	Saudi Arabia	SA	1.96058e+06	19409058	189300	monarchy
45   Singapore	Singapore	SGP	632.6	3396924	66100	republic within Commonwealth
46   Slovakia	Slovakia	SK	48845	5374362	39000	parliamentary democracy
47   Slovenia	Slovenia	SLO	20256	1951443	22600	emerging democracy
48   South Africa	South Africa	RSA	1.21991e+06	41743459	215000	republic
49   South Korea	South Korea	ROK	98480	45482291	590700	republic
50   Spain	Spain	E	504750	39181114	565000	parliamentary monarchy
51   Sri Lanka	Sri Lanka	CL	65610	18553074	65600	republic
52   Sweden	Sweden	S	449964	8900954	177300	constitutional monarchy
53   Switzerland	Switzerland	CH	41290	7207060	158500	federal republic

54	Taiwan	Taiwan	RC	36179	21465881	290500	multiparty democratic regime
55	Tunisia	Tunisia	TN	163610	9019687	37100	republic
56	Turkey	Turkey	TR	780580	62484478	345700	republican parliamentary democracy
57	United Arab Emirates	United Arab Emirates	UAE	75581	3057337	70100	federation with specified powers delegated to the UAE central government and other powers reserved to member emirates
58	United Kingdom	United Kingdom	GB	244820	58489975	1.1384e+06	constitutional monarchy
59	United States	United States	USA	9.37261e+06	266476278	7.2477e+06	federal republic
60	Uruguay	Uruguay	ROU	176220	3238952	24400	republic
61	Vietnam	Vietnam	VN	329560	73976973	97000	Communist state

**Query 3 Results:** This new table reveals that there is other data for these regions relevant for our research. Of course, it is executed in a rather messy format and perhaps worth cleaning up.

**Query 4:** Let's make this table somewhat cleaner by:

- Formatting numbers with ',' as a thousandth's separator
- Shortening government forms to 50 characters

```
In [13]: query = """SELECT t1.country, t2.code, t2.area, t2.population, t3.GDP, SUBSTRING(t4.government,1,50) AS Government
FROM ai_tortoise_indicators AS t1
LEFT JOIN country AS t2 ON t1.country = t2.name
LEFT JOIN economy AS t3 ON t2.code = t3.country
LEFT JOIN politics AS t4 ON t2.code = t4.country
"""

try:
    # Execute the query and store the result in a DataFrame
    df = pd.read_sql_query(query, engine)

    # Apply formatting for thousands and decimal points
    formatted_df = df.style.format({
        'area': '{:,0f}',
        'population': '{:,0f}',
        'GDP': '{:,0f}'
    }, thousands=".")
    # Display the formatted DataFrame
    display(formatted_df)
except Exception as e:
    print(e)
```

	country	code	area	population	GDP	Government
0	Argentina	RA	2.766.890	34.672.997	278.500	republic
1	Armenia	ARM	29.800	3.463.574	9.100	republic
2	Australia	AUS	7.686.850	18.260.863	405.400	federal parliamentary state
3	Austria	A	83.850	8.023.244	152.000	federal republic
4	Bahrain	BRN	620	590.042	7.300	traditional monarchy
5	Belgium	B	30.510	10.170.241	197.000	constitutional monarchy
6	Brazil	BR	8.511.960	162.661.214	976.800	federal republic
7	Canada	CDN	9.976.140	28.820.671	694.000	confederation with parliamentary democracy
8	Chile	RCH	756.950	14.333.258	113.200	republic
9	China	TJ	9.596.960	1.210.004.956	3.500.000	Communist state
10	Colombia	CO	1.138.910	36.813.161	192.500	republic
11	Czech Republic	CZ	78.703	10.321.120	106.200	parliamentary democracy
12	Denmark	DK	43.070	5.249.632	112.800	constitutional monarchy
13	Egypt	ET	1.001.450	63.575.107	171.000	republic
14	Estonia	EW	45.100	1.459.428	12.300	republic
15	Finland	SF	337.030	5.105.230	92.400	republic
16	France	F	547.030	58.317.450	1.173.000	republic
17	Germany	D	356.910	83.536.115	1.452.200	federal republic
18	Greece	GR	131.940	10.538.594	101.700	parliamentary republic
19	Hong Kong	HONX	1.092	7.055.071	307.600	special administrative area in China
20	Hungary	H	93.030	10.002.541	72.500	republic
21	Iceland	IS	103.000	270.292	5.000	republic
22	India	IND	3.287.590	952.107.694	1.408.700	federal republic
23	Indonesia	RI	1.919.440	206.611.600	710.900	republic
24	Ireland	IRL	70.280	3.566.833	54.600	republic
25	Israel	IL	20.770	5.421.995	80.100	republic
26	Italy	I	301.230	57.460.274	1.088.600	republic
27	Japan	J	377.835	125.449.703	2.679.200	constitutional monarchy
28	Kenya	EAK	582.650	28.176.686	36.800	republic
29	Lithuania	LT	65.200	3.646.041	13.300	republic
30	Luxembourg	L	2.586	415.870	10.000	constitutional monarchy

	country	code	area	population	GDP	Government
31	Malaysia	MAL	329.750	19.962.893	193.600	constitutional monarchy
32	Malta	M	320	375.576	4.400	parliamentary democracy
33	Mexico	MEX	1.972.550	95.772.462	721.400	federal republic operating under a centralized gov
34	Morocco	MA	446.550	29.779.156	87.400	constitutional monarchy
35	Netherlands	NL	37.330	15.568.034	301.900	constitutional monarchy
36	New Zealand	NZ	268.680	3.547.983	62.300	parliamentary democracy
37	Nigeria	WAN	923.770	103.912.489	135.900	military government
38	Norway	N	324.220	4.383.807	106.200	constitutional monarchy
39	Pakistan	PK	803.940	129.275.660	274.200	republic
40	Poland	PL	312.683	38.642.565	226.700	democratic state
41	Portugal	P	92.080	9.865.114	116.200	republic
42	Qatar	Q	11.000	547.761	10.700	traditional monarchy
43	Russia	R	17.075.200	148.178.487	796.000	federation
44	Saudi Arabia	SA	1.960.580	19.409.058	189.300	monarchy
45	Singapore	SGP	633	3.396.924	66.100	republic within Commonwealth
46	Slovakia	SK	48.845	5.374.362	39.000	parliamentary democracy
47	Slovenia	SLO	20.256	1.951.443	22.600	emerging democracy
48	South Africa	RSA	1.219.910	41.743.459	215.000	republic
49	South Korea	ROK	98.480	45.482.291	590.700	republic
50	Spain	E	504.750	39.181.114	565.000	parliamentary monarchy
51	Sri Lanka	CL	65.610	18.553.074	65.600	republic
52	Sweden	S	449.964	8.900.954	177.300	constitutional monarchy
53	Switzerland	CH	41.290	7.207.060	158.500	federal republic
54	Taiwan	RC	36.179	21.465.881	290.500	multiparty democratic regime
55	Tunisia	TN	163.610	9.019.687	37.100	republic
56	Turkey	TR	780.580	62.484.478	345.700	republican parliamentary democracy
57	United Arab Emirates	UAE	75.581	3.057.337	70.100	federation with specified powers delegated to the
58	United Kingdom	GB	244.820	58.489.975	1.138.400	constitutional monarchy
59	United States	USA	9.372.610	266.476.278	7.247.700	federal republic
60	Uruguay	ROU	176.220	3.238.952	24.400	republic
61	Vietnam	VN	329.560	73.976.973	97.000	Communist state

**Query 4 Results:** With this cleaned up table, we can easily see that Hong Kong has an area of 1.092 kilometers squared, a population of 7.055.071 people and GDP of 307.600 dollars; similarly, Taiwan has an area of 36.179 kilometers squared, a population of 21.465.881 people, and a GDP of 290.500 dollars. Therefore, we now believe these two regions are worth keeping as independent records in our base view.

**Query 5:** Next, we must develop a view that allows us to execute different queries based on country names or codes and simultaneously reveal basic demographic, economic and political data.

```
In [14]: # The where clause guarantees that there is a connection to the ai_tortoise_indicators table (which means the Hong Kong and Taiwan will be included)
# OR that the country is an official UN country (t2.name is not null if there is a link to the ai_gapminder_country table)
query = """
SELECT t1.code, t1.name, t1.area, t1.population, t2.iso3166_1_alpha2 AS ISO2,
t2.iso3166_1_alpha3 AS ISO3, t2.world_4region, t2.g77_and_oecd_countries AS club, t2.income_groups ,
t3.GDP, t3.GDP/t1.population*100000 AS GDPPC,
SUBSTRING(t4.government,1,50) AS Government,
t5.overall
FROM country AS t1
LEFT JOIN ai_gapminder_country AS t2 ON t1.name = t2.name
LEFT JOIN economy AS t3 ON t1.code = t3.country
LEFT JOIN politics AS t4 ON t1.code = t4.country
LEFT JOIN ai_tortoise_indicators AS t5 ON t1.name = t5.country
where t5.overall IS NOT NULL OR t2.name IS NOT NULL
"""

try:
    # Execute the query and store the result in a DataFrame
    df = pd.read_sql_query(query, engine)

    # Apply formatting for thousands and decimal points
    formatted_df = df.style \
        .format({
            'area': '{:.0f}',
            'population': '{:.0f}',
            'GDP': '{:.0f}',
            'GDPPC': '{:.2f}',
            'overall': '{:.2f}'
        }, thousands=",", decimal=",")

    # Display the formatted DataFrame
    display(formatted_df)
except Exception as e:
    print(e)
```

code	name	area	population	ISO2	ISO3	world_4region	club	income_groups	GDP	GDPPC	Government	overall
0	Austria	83.850	8.023.244	AT	AUT	europe	oecd	high_income	152.000	18.944,96	federal republic	27,70
1	Afghanistan	647.500	22.664.136	AF	AFG	asia	g77	low_income	12.800	564,77	transitional government	nan
2	Antigua and Barbuda	442	65.647	AG	ATG	americas	g77	high_income	425	6.474,02	parliamentary democracy	nan
3	Albania	28.750	3.249.136	AL	ALB	europe	others	upper_middle_income	4.100	1.261,87	emerging democracy	nan
4	Andorra	450	72.766	AD	AND	europe	others	high_income	1.000	13.742,68	parliamentary democracy that retains as its heads	nan
5	Angola	1.246.700	10.342.899	AO	AGO	africa	g77	lower_middle_income	7.400	715,47	transitional government nominally a multiparty dem	nan
6	Armenia	29.800	3.463.574	AM	ARM	europe	others	upper_middle_income	9.100	2.627,34	republic	14,50
7	Australia	7.686.850	18.260.863	AU	AUS	asia	oecd	high_income	405.400	22.200,48	federal parliamentary state	30,90
8	Azerbaijan	86.600	7.676.953	AZ	AZE	europe	others	upper_middle_income	11.500	1.497,99	republic	nan
9	Belgium	30.510	10.170.241	BE	BEL	europe	oecd	high_income	197.000	19.370,24	constitutional monarchy	26,60
10	Bangladesh	144.000	123.062.800	BD	BGD	asia	g77	lower_middle_income	144.500	1.174,20	republic	nan
11	Barbados	430	257.030	BB	BRB	americas	g77	high_income	2.500	9.726,49	parliamentary democracy	nan
12	Benin	112.620	5.709.529	BJ	BEN	africa	g77	low_income	7.600	1.331,11	republic under multiparty democratic rule dropped	nan
13	Burkina Faso	274.200	10.623.323	BF	BFA	africa	g77	low_income	7.400	696,58	parliamentary	nan
14	Bulgaria	110.910	8.612.757	BG	BGR	europe	others	upper_middle_income	43.200	5.015,82	emerging democracy	nan
15	Bhutan	47.000	1.822.625	BT	BTN	asia	g77	lower_middle_income	1.300	713,26	monarchy	nan
16	Burundi	27.830	5.943.057	BI	BDI	africa	g77	low_income	4.000	673,05	republic	nan
17	Bosnia and Herzegovina	51.233	2.656.240	BA	BIH	europe	g77	upper_middle_income	1.000	376,47	emerging democracy	nan
18	Bolivia	1.098.580	7.165.257	BO	BOL	americas	g77	lower_middle_income	20.000	2.791,25	republic	nan
19	Brazil	8.511.960	162.661.214	BR	BRA	americas	g77	upper_middle_income	976.800	6.005,12	federal republic	22,10
20	Bahrain	620	590.042	BH	BHR	asia	g77	high_income	7.300	12.372,00	traditional monarchy	13,50
21	Brunei	5.770	299.939	BN	BRN	asia	g77	high_income	4.600	15.336,45	constitutional sultanate	nan
22	Bahamas	13.940	259.367	BS	BHS	americas	g77	high_income	4.800	18.506,59	commonwealth	nan
23	Belarus	207.600	10.415.973	BY	BLR	europe	others	upper_middle_income	49.200	4.723,51	republic	nan
24	Belize	22.960	219.296	BZ	BLZ	americas	g77	upper_middle_income	575	2.622,03	parliamentary democracy	nan
25	Cuba	110.860	10.951.334	CU	CUB	americas	g77	upper_middle_income	14.700	1.342,30	Communist state	nan
26	Cameroon	475.440	14.261.557	CM	CMR	africa	g77	lower_middle_income	16.500	1.156,96	unitary republic	nan
27	Canada	9.976.140	28.820.671	CA	CAN	americas	oecd	high_income	694.000	24.079,94	confederation with parliamentary democracy	40,30
28	Switzerland	41.290	7.207.060	CH	CHE	europe	oecd	high_income	158.500	21.992,32	federal republic	37,70
29	Sri Lanka	65.610	18.553.074	LK	LKA	asia	g77	upper_middle_income	65.600	3.535,80	republic	10,00
30	Colombia	1.138.910	36.813.161	CO	COL	americas	g77	upper_middle_income	192.500	5.229,11	republic	17,80

	code	name	area	population	ISO2	ISO3	world_4region	club	income_groups	GDP	GDPPC	Government	overall
31	COM	Comoros	2.170	569.237	KM	COM	africa	g77	lower_middle_income	370	649,99	independent republic	nan
32	CR	Costa Rica	51.100	3.463.083	CR	CRI	americas	g77	upper_middle_income	18.400	5.313,18	democratic republic	nan
33	CV	Cape Verde	4.030	449.066	CV	CPV	africa	g77	lower_middle_income	440	979,81	republic	nan
34	CY	Cyprus	9.251	744.609	CY	CYP	europe	others	high_income	7.800	10.475,30	republic	nan
35	CZ	Czech Republic	78.703	10.321.120	CZ	CZE	europe	oecd	high_income	106.200	10.289,58	parliamentary democracy	22,10
36	D	Germany	356.910	83.536.115	DE	DEU	europe	oecd	high_income	1.452.200	17.384,10	federal republic	39,20
37	DJI	Djibouti	22.000	427.642	DJ	DJI	africa	g77	lower_middle_income	500	1.169,20	republic	nan
38	DK	Denmark	43.070	5.249.632	DK	DNK	europe	oecd	high_income	112.800	21.487,22	constitutional monarchy	30,50
39	DOM	Dominican Republic	48.730	8.088.881	DO	DOM	americas	g77	upper_middle_income	26.800	3.313,19	republic	nan
40	DZ	Algeria	2.381.740	29.183.032	DZ	DZA	africa	g77	upper_middle_income	108.700	3.724,77	republic	nan
41	E	Spain	504.750	39.181.114	ES	ESP	europe	oecd	high_income	565.000	14.420,21	parliamentary monarchy	27,70
42	EAK	Kenya	582.650	28.176.686	KE	KEN	africa	g77	lower_middle_income	36.800	1.306,04	republic	8,30
43	EAT	Tanzania	945.090	29.058.470	TZ	TZA	africa	g77	low_income	23.100	794,95	republic	nan
44	EAU	Uganda	236.040	20.158.176	UG	UGA	africa	g77	low_income	16.800	833,41	republic	nan
45	EC	Ecuador	283.560	11.466.291	EC	ECU	americas	g77	upper_middle_income	44.600	3.889,66	republic	nan
46	ER	Eritrea	121.320	3.427.883	ER	ERI	africa	g77	low_income	2.000	583,45	transitional government	nan
47	ES	El Salvador	21.040	5.828.987	SV	SLV	americas	g77	lower_middle_income	11.400	1.955,74	republic	nan
48	ET	Egypt	1.001.450	63.575.107	EG	EGY	africa	g77	lower_middle_income	171.000	2.689,73	republic	16,90
49	ETH	Ethiopia	1.127.130	57.171.662	ET	ETH	africa	g77	low_income	24.200	423,29	federal republic	nan
50	EW	Estonia	45.100	1.459.428	EE	EST	europe	others	high_income	12.300	8.427,96	republic	26,00
51	F	France	547.030	58.317.450	FR	FRA	europe	oecd	high_income	1.173.000	20.114,05	republic	32,80
52	FJI	Fiji	18.270	782.381	FJ	FJI	asia	g77	upper_middle_income	4.700	6.007,30	republic	nan
53	FL	Liechtenstein	160	31.122	LI	LIE	europe	others	high_income	630	20.242,91	hereditary constitutional monarchy	nan
54	G	Gabon	267.670	1.172.798	GA	GAB	africa	g77	upper_middle_income	6.000	5.115,97	republic	nan
55	GB	United Kingdom	244.820	58.489.975	GB	GBR	europe	oecd	high_income	1.138.400	19.463,16	constitutional monarchy	41,80
56	GCA	Guatemala	108.890	11.277.614	GT	GTM	americas	g77	upper_middle_income	36.700	3.254,23	republic	nan
57	GE	Georgia	69.700	5.219.810	GE	GEO	europe	others	upper_middle_income	6.200	1.187,78	republic	nan
58	GH	Ghana	238.540	17.698.271	GH	GHA	africa	g77	lower_middle_income	25.100	1.418,22	constitutional democracy	nan
59	GNB	Guinea-Bissau	36.120	1.151.330	GW	GNB	africa	g77	low_income	1.000	868,56	republic	nan
60	GQ	Equatorial Guinea	28.050	431.282	GQ	GNQ	africa	g77	upper_middle_income	325	753,57	republic in transition to multiparty democracy	nan
61	GR	Greece	131.940	10.538.594	GR	GRC	europe	oecd	high_income	101.700	9.650,24	parliamentary republic	18,30

code	name	area	population	ISO2	ISO3	world_4region	club	income_groups	GDP	GDPPC	Government	overall
62	GUY	Guyana	214.970	712.091	GY	GUY	americas	g77	upper_middle_income	1.600	2.246,90	republic nan
63	H	Hungary	93.030	10.002.541	HU	HUN	europe	oecd	high_income	72.500	7.248,16	republic 20,70
64	HCA	Honduras	112.090	5.605.193	HN	HND	americas	g77	lower_middle_income	10.800	1.926,78	republic nan
65	HONX	Hong Kong	1.092	7.055.071	None	None	None	None	None	307.600	43.599,84	special administrative area in China 22,50
66	HR	Croatia	56.538	5.004.112	HR	HRV	europe	others	high_income	20.100	4.016,70	parliamentary democracy nan
67	I	Italy	301.230	57.460.274	IT	ITA	europe	oecd	high_income	1.088.600	18.945,26	republic 26,50
68	IL	Israel	20.770	5.421.995	IL	ISR	asia	oecd	high_income	80.100	14.773,16	republic 40,00
69	IND	India	3.287.590	952.107.694	IN	IND	asia	g77	lower_middle_income	1.408.700	1.479,56	federal republic 31,40
70	IR	Iran	1.648.000	66.094.264	IR	IRN	asia	g77	upper_middle_income	323.500	4.894,52	theocratic republic nan
71	IRL	Ireland	70.280	3.566.833	IE	IRL	europe	oecd	high_income	54.600	15.307,70	republic 28,80
72	IRQ	Iraq	437.072	21.422.292	IQ	IRQ	asia	g77	upper_middle_income	41.100	1.918,56	republic nan
73	IS	Iceland	103.000	270.292	IS	ISL	europe	oecd	high_income	5.000	18.498,51	republic 20,60
74	J	Japan	377.835	125.449.703	JP	JPN	asia	oecd	high_income	2.679.200	21.356,77	constitutional monarchy 33,90
75	JA	Jamaica	10.991	2.595.275	JM	JAM	americas	g77	upper_middle_income	8.200	3.159,59	parliamentary democracy nan
76	JOR	Jordan	89.213	4.212.152	JO	JOR	asia	g77	upper_middle_income	19.300	4.581,98	constitutional monarchy nan
77	K	Cambodia	181.040	10.861.218	KH	KHM	asia	g77	lower_middle_income	7.000	644,49	multiparty liberal democracy under a constitutiona nan
78	KIR	Kiribati	717	80.919	KI	KIR	asia	others	lower_middle_income	68	840,35	republic nan
79	KWT	Kuwait	17.820	1.950.047	KW	KWT	asia	g77	high_income	30.800	15.794,49	nominal constitutional monarchy nan
80	L	Luxembourg	2.586	415.870	LU	LUX	europe	oecd	high_income	10.000	24.045,98	constitutional monarchy 29,20
81	LAR	Libya	1.759.540	5.445.436	LY	LYB	africa	g77	upper_middle_income	32.900	6.041,76	Jamahiriya in theory nan
82	LB	Liberia	111.370	2.109.789	LR	LBL	africa	g77	low_income	2.300	1.090,16	republic nan
83	LS	Lesotho	30.350	1.970.781	LS	LSO	africa	g77	lower_middle_income	2.800	1.420,76	modified constitutional monarchy nan
84	LT	Lithuania	65.200	3.646.041	LT	LTU	europe	others	high_income	13.300	3.647,79	republic 19,70
85	LV	Latvia	64.100	2.468.982	LV	LVA	europe	others	high_income	14.700	5.953,87	republic nan
86	M	Malta	320	375.576	MT	MLT	europe	others	high_income	4.400	11.715,34	parliamentary democracy 22,40
87	MA	Morocco	446.550	29.779.156	MA	MAR	africa	g77	lower_middle_income	87.400	2.934,94	constitutional monarchy 13,60
88	MAL	Malaysia	329.750	19.962.893	MY	MYS	asia	g77	upper_middle_income	193.600	9.697,99	constitutional monarchy 19,60
89	MC	Monaco	2	31.719	MC	MCO	europe	others	high_income	788	24.843,15	constitutional monarchy nan
90	MD	Moldova	33.700	4.463.847	MD	MDA	europe	others	lower_middle_income	10.400	2.329,83	republic nan
91	MEX	Mexico	1.972.550	95.772.462	MX	MEX	americas	oecd	upper_middle_income	721.400	7.532,44	federal republic operating under a centralized gov 16,90
92	MH	Marshall Islands	181	58.363	MH	MHL	asia	g77	upper_middle_income	94	1.610,61	constitutional government in free association with nan

	code	name	area	population	ISO2	ISO3	world_4region	club	income_groups	GDP	GDPPC	Government	overall
93	MNE	Montenegro	14.026	672.180	ME	MNE	europe	others	upper_middle_income	4.515	6.716,95	parliamentary democracy	nan
94	MNG	Mongolia	1.565.000	2.496.617	MN	MNG	asia	g77	lower_middle_income	4.900	1.962,66	republic	nan
95	MOC	Mozambique	801.590	17.877.927	MZ	MOZ	africa	g77	low_income	12.200	682,41	republic	nan
96	MS	Mauritius	2.040	1.140.256	MU	MUS	africa	g77	upper_middle_income	10.900	9.559,26	parliamentary democracy	nan
97	MV	Maldives	300	270.758	MV	MDV	asia	g77	upper_middle_income	390	1.440,40	republic	nan
98	MW	Malawi	118.480	9.452.844	MW	MWI	africa	g77	low_income	6.900	729,94	multiparty democracy	nan
99	MYA	Myanmar	678.500	45.975.625	MM	MMR	asia	g77	lower_middle_income	47.000	1.022,28	military regime	nan
100	N	Norway	324.220	4.383.807	NO	NOR	europe	oecd	high_income	106.200	24.225,52	constitutional monarchy	26,40
101	NAM	Namibia	825.418	1.677.243	NA	NAM	africa	g77	upper_middle_income	5.800	3.458,06	republic	nan
102	NAU	Nauru	21	10.273	NR	NRU	asia	others	upper_middle_income	100	9.734,25	republic	nan
103	NEP	Nepal	140.800	22.094.033	NP	NPL	asia	g77	low_income	25.200	1.140,58	parliamentary democracy as of 12 May 1991	nan
104	NIC	Nicaragua	129.494	4.272.352	NI	NIC	americas	g77	lower_middle_income	7.100	1.661,85	republic	nan
105	NL	Netherlands	37.330	15.568.034	NL	NLD	europe	oecd	high_income	301.900	19.392,30	constitutional monarchy	34,50
106	NOK	North Korea	120.540	23.904.124	KP	PRK	asia	g77	low_income	21.500	899,43	Communist state	nan
107	NZ	New Zealand	268.680	3.547.983	NZ	NZL	asia	oecd	high_income	62.300	17.559,27	parliamentary democracy	21,60
108	OM	Oman	212.460	2.186.548	OM	OMN	asia	g77	high_income	19.100	8.735,23	monarchy	nan
109	P	Portugal	92.080	9.865.114	PT	PRT	europe	oecd	high_income	116.200	11.778,88	republic	23,70
110	PA	Panama	78.200	2.655.094	PA	PAN	americas	g77	high_income	13.600	5.122,23	constitutional republic	nan
111	PAL	Palau	458	16.952	PW	PLW	asia	others	high_income	82	4.825,39	constitutional government in free association with	nan
112	PE	Peru	1.285.220	24.523.408	PE	PER	americas	g77	upper_middle_income	87.000	3.547,63	republic	nan
113	PK	Pakistan	803.940	129.275.660	PK	PAK	asia	g77	lower_middle_income	274.200	2.121,05	republic	10,10
114	PL	Poland	312.683	38.642.565	PL	POL	europe	oecd	high_income	226.700	5.866,59	democratic state	24,80
115	PNG	Papua New Guinea	461.690	4.394.537	PG	PNG	asia	g77	lower_middle_income	10.200	2.321,06	parliamentary democracy	nan
116	PY	Paraguay	406.750	5.504.146	PY	PRY	americas	g77	upper_middle_income	17.000	3.088,58	republic	nan
117	Q	Qatar	11.000	547.761	QA	QAT	asia	g77	high_income	10.700	19.534,07	traditional monarchy	19,80
118	R	Russia	17.075.200	148.178.487	RU	RUS	europe	others	upper_middle_income	796.000	5.371,90	federation	23,70
119	RA	Argentina	2.766.890	34.672.997	AR	ARG	americas	g77	upper_middle_income	278.500	8.032,19	republic	17,50
120	RB	Botswana	600.370	1.477.630	BW	BWA	africa	g77	upper_middle_income	4.500	3.045,42	parliamentary republic	nan
121	RC	Taiwan	36.179	21.465.881	None	None	None	None	None	290.500	13.533,10	multiparty democratic regime	25,40
122	RCA	Central African Republic	622.980	3.274.426	CF	CAF	africa	g77	low_income	2.500	763,49	republic	nan
123	RCH	Chile	756.950	14.333.258	CL	CHL	americas	g77	high_income	113.200	7.897,72	republic	20,20

	code	name	area	population	ISO2	ISO3	world_4region	club	income_groups	GDP	GDPPC	Government	overall
124	RG	Guinea	245.860	7.411.981	GN	GIN	africa	g77	low_income	6.500	876,96	republic	nan
125	RH	Haiti	27.750	6.731.539	HT	HTI	americas	g77	low_income	6.500	965,60	republic	nan
126	RI	Indonesia	1.919.440	206.611.600	ID	IDN	asia	g77	lower_middle_income	710.900	3.440,76	republic	18,20
127	RIM	Mauritania	1.030.700	2.336.048	MR	MRT	africa	g77	lower_middle_income	2.800	1.198,61	republic	nan
128	RL	Lebanon	10.400	3.776.317	LB	LBN	asia	g77	upper_middle_income	18.300	4.845,99	republic	nan
129	RM	Madagascar	587.041	13.670.507	MG	MDG	africa	g77	low_income	11.400	833,91	republic	nan
130	RMM	Mali	1.240.000	9.653.261	ML	MLI	africa	g77	low_income	5.400	559,40	republic	nan
131	RN	Niger	1.267.000	9.113.001	NE	NER	africa	g77	low_income	5.500	603,53	republic	nan
132	RO	Romania	237.500	21.657.162	RO	ROU	europe	others	upper_middle_income	105.700	4.880,60	republic	nan
133	ROK	South Korea	98.480	45.482.291	KR	KOR	asia	oecd	high_income	590.700	12.987,47	republic	40,30
134	ROU	Uruguay	176.220	3.238.952	UY	URY	americas	g77	high_income	24.400	7.533,30	republic	16,30
135	RP	Philippines	299.764	74.480.848	PH	PHL	asia	g77	lower_middle_income	179.700	2.412,70	republic	nan
136	RSA	South Africa	1.219.910	41.743.459	ZA	ZAF	africa	g77	upper_middle_income	215.000	5.150,51	republic	14,10
137	RSM	San Marino	60	24.521	SM	SMR	europe	others	high_income	380	15.496,92	republic	nan
138	RT	Togo	56.790	4.570.530	TG	TGO	africa	g77	low_income	4.100	897,05	republic under transition to multiparty democratic	nan
139	RWA	Rwanda	26.340	6.853.359	RW	RWA	africa	g77	low_income	3.800	554,47	republic	nan
140	S	Sweden	449.964	8.900.954	SE	SWE	europe	oecd	high_income	177.300	19.919,21	constitutional monarchy	30,30
141	SA	Saudi Arabia	1.960.580	19.409.058	SA	SAU	asia	g77	high_income	189.300	9.753,18	monarchy	23,30
142	SF	Finland	337.030	5.105.230	FI	FIN	europe	oecd	high_income	92.400	18.099,09	republic	34,90
143	SGP	Singapore	633	3.396.924	SG	SGP	asia	g77	high_income	66.100	19.458,78	republic within Commonwealth	49,70
144	SK	Slovakia	48.845	5.374.362	SK	SVK	europe	oecd	high_income	39.000	7.256,68	parliamentary democracy	17,10
145	SLB	Solomon Islands	28.450	412.902	SB	SLB	asia	g77	lower_middle_income	1.000	2.421,88	parliamentary democracy	nan
146	SLO	Slovenia	20.256	1.951.443	SI	SVN	europe	others	high_income	22.600	11.581,17	emerging democracy	21,50
147	SME	Suriname	163.270	436.418	SR	SUR	americas	g77	upper_middle_income	1.300	2.978,80	republic	nan
148	SN	Senegal	196.190	9.092.749	SN	SEN	africa	g77	lower_middle_income	14.500	1.594,68	republic under multiparty democratic rule	nan
149	SP	Somalia	637.660	9.639.151	SO	SOM	africa	g77	low_income	3.600	373,48	none	nan
150	SRB	Serbia	77.474	7.379.339	RS	SRB	europe	others	upper_middle_income	52.180	7.071,09	parliamentary democracy	nan
151	STP	Sao Tome and Principe	1.001	144.128	ST	STP	africa	g77	lower_middle_income	138	957,48	republic	nan
152	SUD	Sudan	2.505.810	31.547.543	SD	SDN	africa	g77	lower_middle_income	25.000	792,45	transitional previously ruling military junta	nan
153	SY	Seychelles	455	77.575	SC	SYC	africa	g77	high_income	430	5.543,02	republic	nan
154	SYR	Syria	185.180	15.608.648	SY	SYR	asia	g77	low_income	91.200	5.842,91	republic under military regime since March 1963	nan

code	name	area	population	ISO2	ISO3	world_4region	club	income_groups	GDP	GDPPC	Government	overall
155	TAD	Tajikistan	143.100	5.916.373	TJ	TJK	asia	others	low_income	6.400	1.081,74	republic nan
156	TCH	Chad	1.284.000	6.976.845	TD	TCD	africa	g77	low_income	3.300	472,99	republic nan
157	THA	Thailand	514.000	58.851.357	TH	THA	asia	g77	upper_middle_income	416.700	7.080,55	constitutional monarchy nan
158	TJ	China	9.596.960	1.210.004.956	CN	CHN	asia	g77	upper_middle_income	3.500.000	2.892,55	Communist state 61,50
159	TL	Timor-Leste	15.007	1.131.612	TL	TLS	asia	g77	lower_middle_income	489	432,13	parliamentary democracy nan
160	TM	Turkmenistan	488.100	4.149.283	TM	TKM	asia	g77	upper_middle_income	11.500	2.771,56	republic nan
161	TN	Tunisia	163.610	9.019.687	TN	TUN	africa	g77	lower_middle_income	37.100	4.113,22	republic 13,70
162	TO	Tonga	748	106.466	TO	TON	asia	g77	upper_middle_income	228	2.141,53	hereditary constitutional monarchy nan
163	TR	Turkey	780.580	62.484.478	TR	TUR	europe	oecd	upper_middle_income	345.700	5.532,57	republican parliamentary democracy 20,60
164	TT	Trinidad and Tobago	5.130	1.272.385	TT	TTO	americas	g77	high_income	16.200	12.732,00	parliamentary democracy nan
165	TUV	Tuvalu	26	10.146	TV	TUV	asia	others	upper_middle_income	8	768,78	democracy nan
166	UA	Ukraine	603.700	50.864.009	UA	UKR	europe	others	lower_middle_income	174.600	3.432,68	republic nan
167	UAE	United Arab Emirates	75.581	3.057.337	AE	ARE	asia	g77	high_income	70.100	22.928,45	federation with specified powers delegated to the 23,90
168	USA	United States	9.372.610	266.476.278	US	USA	americas	oecd	high_income	7.247.700	27.198,29	federal republic 100,00
169	UZB	Uzbekistan	447.400	23.418.381	UZ	UZB	asia	others	lower_middle_income	54.700	2.335,77	republic nan
170	V	Holy See	0	840	VA	VAT	europe	others	None	nan	nan	monarchical sacerdotal state nan
171	VN	Vietnam	329.560	73.976.973	VN	VNM	asia	g77	lower_middle_income	97.000	1.311,22	Communist state 18,00
172	VU	Vanuatu	14.760	177.504	VU	VUT	asia	g77	lower_middle_income	210	1.183,07	republic nan
173	WAG	Gambia	11.300	1.204.984	GM	GMB	africa	g77	low_income	1.100	912,88	republic under multiparty democratic rule nan
174	WAL	Sierra Leone	71.740	4.793.121	SL	SLE	africa	g77	low_income	4.400	917,98	constitutional democracy nan
175	WAN	Nigeria	923.770	103.912.489	NG	NGA	africa	g77	lower_middle_income	135.900	1.307,83	military government 9,30
176	WD	Dominica	746	82.926	DM	DMA	americas	g77	upper_middle_income	200	2.411,79	parliamentary democracy nan
177	WG	Grenada	344	94.961	GD	GRD	americas	g77	upper_middle_income	284	2.990,70	parliamentary democracy nan
178	WS	Samoa	2.860	214.384	WS	WSM	asia	g77	upper_middle_income	415	1.935,78	constitutional monarchy under native chief nan
179	YE	Yemen	527.970	13.483.178	YE	YEM	asia	g77	low_income	37.100	2.751,58	republic nan
180	YV	Venezuela	912.050	21.983.188	VE	VEN	americas	g77	upper_middle_income	195.500	8.893,16	republic nan
181	Z	Zambia	752.610	9.159.072	ZM	ZMB	africa	g77	lower_middle_income	8.900	971,71	republic nan
182	ZW	Zimbabwe	390.580	11.271.314	ZW	ZWE	africa	g77	lower_middle_income	18.100	1.605,85	parliamentary democracy nan

**Query 5 Results:** The result is a clean table that includes country names, codes, and key statistics such as the area, population, region, club, income grouping, GDP, Government, and overall scores ascribed by Tortoise Media (more on this later).

**Query 6:** Since everything looks well-organized, we now want to create these results as a view. If the view already exists, we delete it and recreate it.

```
In [15]: view_name = 'basic_country_data'

query = f"""CREATE VIEW {view_name} AS
SELECT t1.code, t1.name, t1.area, t1.population,
t2.iso3166_1_alpha2 AS ISO2, t2.iso3166_1_alpha3 AS ISO3, t2.world_4region, t2.g77_and_oecd_countries AS club, t2.income_groups,
t3.GDP, t3.GDP/t1.population*1000000 AS GDPPC,
SUBSTRING(t4.government,1,50) AS Government,
t5.overall
FROM country AS t1
LEFT JOIN ai_gapminder_country AS t2 ON t1.name = t2.name
LEFT JOIN economy AS t3 ON t1.code = t3.country
LEFT JOIN politics AS t4 ON t1.code = t4.country
LEFT JOIN ai_tortoise_indicators AS t5 ON t1.name = t5.country
WHERE t5.overall IS NOT NULL OR t2.name IS NOT NULL;
"""

try:
    with engine.connect() as conn:
        drop_view_query = f"DROP VIEW IF EXISTS {view_name}"
        conn.execute(text(drop_view_query))
        conn.execute(text(query))
except Exception as e:
    print(e)
```

**Query 6 Results:** The output is error-free, confirming that our code was inputted correctly. However, even if correct, this does not necessarily mean it contains the data we need.

**Query 7:** To confirm that the created view contains the expected data, we now perform a simple query of the view.

```
In [16]: query = f"SELECT * FROM {view_name}"

try:
    # Execute the query and store the result in a DataFrame
    df = pd.read_sql_query(query, engine)

    # Apply formatting for thousands and decimal points
    formatted_df = df.style \
        .format({
            'area': '{:.0f}',
            'population': '{:.0f}',
            'GDP': '{:.0f}',
            'GDPPC': '{:.2f}',
            'overall': '{:.2f}'
        }, thousands=",", decimal=",")

    # Display the formatted DataFrame
    display(formatted_df)
except Exception as e:
    print(e)
```

code	name	area	population	ISO2	ISO3	world_4region	club	income_groups	GDP	GDPPC	Government	overall
0	Austria	83.850	8.023.244	AT	AUT	europe	oecd	high_income	152.000	18.944,96	federal republic	27,70
1	Afghanistan	647.500	22.664.136	AF	AFG	asia	g77	low_income	12.800	564,77	transitional government	nan
2	Antigua and Barbuda	442	65.647	AG	ATG	americas	g77	high_income	425	6.474,02	parliamentary democracy	nan
3	Albania	28.750	3.249.136	AL	ALB	europe	others	upper_middle_income	4.100	1.261,87	emerging democracy	nan
4	Andorra	450	72.766	AD	AND	europe	others	high_income	1.000	13.742,68	parliamentary democracy that retains as its heads	nan
5	Angola	1.246.700	10.342.899	AO	AGO	africa	g77	lower_middle_income	7.400	715,47	transitional government nominally a multiparty dem	nan
6	Armenia	29.800	3.463.574	AM	ARM	europe	others	upper_middle_income	9.100	2.627,34	republic	14,50
7	Australia	7.686.850	18.260.863	AU	AUS	asia	oecd	high_income	405.400	22.200,48	federal parliamentary state	30,90
8	Azerbaijan	86.600	7.676.953	AZ	AZE	europe	others	upper_middle_income	11.500	1.497,99	republic	nan
9	Belgium	30.510	10.170.241	BE	BEL	europe	oecd	high_income	197.000	19.370,24	constitutional monarchy	26,60
10	Bangladesh	144.000	123.062.800	BD	BGD	asia	g77	lower_middle_income	144.500	1.174,20	republic	nan
11	Barbados	430	257.030	BB	BRB	americas	g77	high_income	2.500	9.726,49	parliamentary democracy	nan
12	Benin	112.620	5.709.529	BJ	BEN	africa	g77	low_income	7.600	1.331,11	republic under multiparty democratic rule dropped	nan
13	Burkina Faso	274.200	10.623.323	BF	BFA	africa	g77	low_income	7.400	696,58	parliamentary	nan
14	Bulgaria	110.910	8.612.757	BG	BGR	europe	others	upper_middle_income	43.200	5.015,82	emerging democracy	nan
15	Bhutan	47.000	1.822.625	BT	BTN	asia	g77	lower_middle_income	1.300	713,26	monarchy	nan
16	Burundi	27.830	5.943.057	BI	BDI	africa	g77	low_income	4.000	673,05	republic	nan
17	Bosnia and Herzegovina	51.233	2.656.240	BA	BIH	europe	g77	upper_middle_income	1.000	376,47	emerging democracy	nan
18	Bolivia	1.098.580	7.165.257	BO	BOL	americas	g77	lower_middle_income	20.000	2.791,25	republic	nan
19	Brazil	8.511.960	162.661.214	BR	BRA	americas	g77	upper_middle_income	976.800	6.005,12	federal republic	22,10
20	Bahrain	620	590.042	BH	BHR	asia	g77	high_income	7.300	12.372,00	traditional monarchy	13,50
21	Brunei	5.770	299.939	BN	BRN	asia	g77	high_income	4.600	15.336,45	constitutional sultanate	nan
22	Bahamas	13.940	259.367	BS	BHS	americas	g77	high_income	4.800	18.506,59	commonwealth	nan
23	Belarus	207.600	10.415.973	BY	BLR	europe	others	upper_middle_income	49.200	4.723,51	republic	nan
24	Belize	22.960	219.296	BZ	BLZ	americas	g77	upper_middle_income	575	2.622,03	parliamentary democracy	nan
25	Cuba	110.860	10.951.334	CU	CUB	americas	g77	upper_middle_income	14.700	1.342,30	Communist state	nan
26	Cameroon	475.440	14.261.557	CM	CMR	africa	g77	lower_middle_income	16.500	1.156,96	unitary republic	nan
27	Canada	9.976.140	28.820.671	CA	CAN	americas	oecd	high_income	694.000	24.079,94	confederation with parliamentary democracy	40,30
28	Switzerland	41.290	7.207.060	CH	CHE	europe	oecd	high_income	158.500	21.992,32	federal republic	37,70
29	Sri Lanka	65.610	18.553.074	LK	LKA	asia	g77	upper_middle_income	65.600	3.535,80	republic	10,00
30	Colombia	1.138.910	36.813.161	CO	COL	americas	g77	upper_middle_income	192.500	5.229,11	republic	17,80

	code	name	area	population	ISO2	ISO3	world_4region	club	income_groups	GDP	GDPPC	Government	overall
31	COM	Comoros	2.170	569.237	KM	COM	africa	g77	lower_middle_income	370	649,99	independent republic	nan
32	CR	Costa Rica	51.100	3.463.083	CR	CRI	americas	g77	upper_middle_income	18.400	5.313,18	democratic republic	nan
33	CV	Cape Verde	4.030	449.066	CV	CPV	africa	g77	lower_middle_income	440	979,81	republic	nan
34	CY	Cyprus	9.251	744.609	CY	CYP	europe	others	high_income	7.800	10.475,30	republic	nan
35	CZ	Czech Republic	78.703	10.321.120	CZ	CZE	europe	oecd	high_income	106.200	10.289,58	parliamentary democracy	22,10
36	D	Germany	356.910	83.536.115	DE	DEU	europe	oecd	high_income	1.452.200	17.384,10	federal republic	39,20
37	DJI	Djibouti	22.000	427.642	DJ	DJI	africa	g77	lower_middle_income	500	1.169,20	republic	nan
38	DK	Denmark	43.070	5.249.632	DK	DNK	europe	oecd	high_income	112.800	21.487,22	constitutional monarchy	30,50
39	DOM	Dominican Republic	48.730	8.088.881	DO	DOM	americas	g77	upper_middle_income	26.800	3.313,19	republic	nan
40	DZ	Algeria	2.381.740	29.183.032	DZ	DZA	africa	g77	upper_middle_income	108.700	3.724,77	republic	nan
41	E	Spain	504.750	39.181.114	ES	ESP	europe	oecd	high_income	565.000	14.420,21	parliamentary monarchy	27,70
42	EAK	Kenya	582.650	28.176.686	KE	KEN	africa	g77	lower_middle_income	36.800	1.306,04	republic	8,30
43	EAT	Tanzania	945.090	29.058.470	TZ	TZA	africa	g77	low_income	23.100	794,95	republic	nan
44	EAU	Uganda	236.040	20.158.176	UG	UGA	africa	g77	low_income	16.800	833,41	republic	nan
45	EC	Ecuador	283.560	11.466.291	EC	ECU	americas	g77	upper_middle_income	44.600	3.889,66	republic	nan
46	ER	Eritrea	121.320	3.427.883	ER	ERI	africa	g77	low_income	2.000	583,45	transitional government	nan
47	ES	El Salvador	21.040	5.828.987	SV	SLV	americas	g77	lower_middle_income	11.400	1.955,74	republic	nan
48	ET	Egypt	1.001.450	63.575.107	EG	EGY	africa	g77	lower_middle_income	171.000	2.689,73	republic	16,90
49	ETH	Ethiopia	1.127.130	57.171.662	ET	ETH	africa	g77	low_income	24.200	423,29	federal republic	nan
50	EW	Estonia	45.100	1.459.428	EE	EST	europe	others	high_income	12.300	8.427,96	republic	26,00
51	F	France	547.030	58.317.450	FR	FRA	europe	oecd	high_income	1.173.000	20.114,05	republic	32,80
52	FJI	Fiji	18.270	782.381	FJ	FJI	asia	g77	upper_middle_income	4.700	6.007,30	republic	nan
53	FL	Liechtenstein	160	31.122	LI	LIE	europe	others	high_income	630	20.242,91	hereditary constitutional monarchy	nan
54	G	Gabon	267.670	1.172.798	GA	GAB	africa	g77	upper_middle_income	6.000	5.115,97	republic	nan
55	GB	United Kingdom	244.820	58.489.975	GB	GBR	europe	oecd	high_income	1.138.400	19.463,16	constitutional monarchy	41,80
56	GCA	Guatemala	108.890	11.277.614	GT	GTM	americas	g77	upper_middle_income	36.700	3.254,23	republic	nan
57	GE	Georgia	69.700	5.219.810	GE	GEO	europe	others	upper_middle_income	6.200	1.187,78	republic	nan
58	GH	Ghana	238.540	17.698.271	GH	GHA	africa	g77	lower_middle_income	25.100	1.418,22	constitutional democracy	nan
59	GNB	Guinea-Bissau	36.120	1.151.330	GW	GNB	africa	g77	low_income	1.000	868,56	republic	nan
60	GQ	Equatorial Guinea	28.050	431.282	GQ	GNQ	africa	g77	upper_middle_income	325	753,57	republic in transition to multiparty democracy	nan
61	GR	Greece	131.940	10.538.594	GR	GRC	europe	oecd	high_income	101.700	9.650,24	parliamentary republic	18,30

code	name	area	population	ISO2	ISO3	world_4region	club	income_groups	GDP	GDPPC	Government	overall
62	GUY	Guyana	214.970	712.091	GY	GUY	americas	g77	upper_middle_income	1.600	2.246,90	republic nan
63	H	Hungary	93.030	10.002.541	HU	HUN	europe	oecd	high_income	72.500	7.248,16	republic 20,70
64	HCA	Honduras	112.090	5.605.193	HN	HND	americas	g77	lower_middle_income	10.800	1.926,78	republic nan
65	HONX	Hong Kong	1.092	7.055.071	None	None	None	None	None	307.600	43.599,84	special administrative area in China 22,50
66	HR	Croatia	56.538	5.004.112	HR	HRV	europe	others	high_income	20.100	4.016,70	parliamentary democracy nan
67	I	Italy	301.230	57.460.274	IT	ITA	europe	oecd	high_income	1.088.600	18.945,26	republic 26,50
68	IL	Israel	20.770	5.421.995	IL	ISR	asia	oecd	high_income	80.100	14.773,16	republic 40,00
69	IND	India	3.287.590	952.107.694	IN	IND	asia	g77	lower_middle_income	1.408.700	1.479,56	federal republic 31,40
70	IR	Iran	1.648.000	66.094.264	IR	IRN	asia	g77	upper_middle_income	323.500	4.894,52	theocratic republic nan
71	IRL	Ireland	70.280	3.566.833	IE	IRL	europe	oecd	high_income	54.600	15.307,70	republic 28,80
72	IRQ	Iraq	437.072	21.422.292	IQ	IRQ	asia	g77	upper_middle_income	41.100	1.918,56	republic nan
73	IS	Iceland	103.000	270.292	IS	ISL	europe	oecd	high_income	5.000	18.498,51	republic 20,60
74	J	Japan	377.835	125.449.703	JP	JPN	asia	oecd	high_income	2.679.200	21.356,77	constitutional monarchy 33,90
75	JA	Jamaica	10.991	2.595.275	JM	JAM	americas	g77	upper_middle_income	8.200	3.159,59	parliamentary democracy nan
76	JOR	Jordan	89.213	4.212.152	JO	JOR	asia	g77	upper_middle_income	19.300	4.581,98	constitutional monarchy nan
77	K	Cambodia	181.040	10.861.218	KH	KHM	asia	g77	lower_middle_income	7.000	644,49	multiparty liberal democracy under a constitutiona nan
78	KIR	Kiribati	717	80.919	KI	KIR	asia	others	lower_middle_income	68	840,35	republic nan
79	KWT	Kuwait	17.820	1.950.047	KW	KWT	asia	g77	high_income	30.800	15.794,49	nominal constitutional monarchy nan
80	L	Luxembourg	2.586	415.870	LU	LUX	europe	oecd	high_income	10.000	24.045,98	constitutional monarchy 29,20
81	LAR	Libya	1.759.540	5.445.436	LY	LYB	africa	g77	upper_middle_income	32.900	6.041,76	Jamahiriya in theory nan
82	LB	Liberia	111.370	2.109.789	LR	LBL	africa	g77	low_income	2.300	1.090,16	republic nan
83	LS	Lesotho	30.350	1.970.781	LS	LSO	africa	g77	lower_middle_income	2.800	1.420,76	modified constitutional monarchy nan
84	LT	Lithuania	65.200	3.646.041	LT	LTU	europe	others	high_income	13.300	3.647,79	republic 19,70
85	LV	Latvia	64.100	2.468.982	LV	LVA	europe	others	high_income	14.700	5.953,87	republic nan
86	M	Malta	320	375.576	MT	MLT	europe	others	high_income	4.400	11.715,34	parliamentary democracy 22,40
87	MA	Morocco	446.550	29.779.156	MA	MAR	africa	g77	lower_middle_income	87.400	2.934,94	constitutional monarchy 13,60
88	MAL	Malaysia	329.750	19.962.893	MY	MYS	asia	g77	upper_middle_income	193.600	9.697,99	constitutional monarchy 19,60
89	MC	Monaco	2	31.719	MC	MCO	europe	others	high_income	788	24.843,15	constitutional monarchy nan
90	MD	Moldova	33.700	4.463.847	MD	MDA	europe	others	lower_middle_income	10.400	2.329,83	republic nan
91	MEX	Mexico	1.972.550	95.772.462	MX	MEX	americas	oecd	upper_middle_income	721.400	7.532,44	federal republic operating under a centralized gov 16,90
92	MH	Marshall Islands	181	58.363	MH	MHL	asia	g77	upper_middle_income	94	1.610,61	constitutional government in free association with nan

	code	name	area	population	ISO2	ISO3	world_4region	club	income_groups	GDP	GDPPC	Government	overall
93	MNE	Montenegro	14.026	672.180	ME	MNE	europe	others	upper_middle_income	4.515	6.716,95	parliamentary democracy	nan
94	MNG	Mongolia	1.565.000	2.496.617	MN	MNG	asia	g77	lower_middle_income	4.900	1.962,66	republic	nan
95	MOC	Mozambique	801.590	17.877.927	MZ	MOZ	africa	g77	low_income	12.200	682,41	republic	nan
96	MS	Mauritius	2.040	1.140.256	MU	MUS	africa	g77	upper_middle_income	10.900	9.559,26	parliamentary democracy	nan
97	MV	Maldives	300	270.758	MV	MDV	asia	g77	upper_middle_income	390	1.440,40	republic	nan
98	MW	Malawi	118.480	9.452.844	MW	MWI	africa	g77	low_income	6.900	729,94	multiparty democracy	nan
99	MYA	Myanmar	678.500	45.975.625	MM	MMR	asia	g77	lower_middle_income	47.000	1.022,28	military regime	nan
100	N	Norway	324.220	4.383.807	NO	NOR	europe	oecd	high_income	106.200	24.225,52	constitutional monarchy	26,40
101	NAM	Namibia	825.418	1.677.243	NA	NAM	africa	g77	upper_middle_income	5.800	3.458,06	republic	nan
102	NAU	Nauru	21	10.273	NR	NRU	asia	others	upper_middle_income	100	9.734,25	republic	nan
103	NEP	Nepal	140.800	22.094.033	NP	NPL	asia	g77	low_income	25.200	1.140,58	parliamentary democracy as of 12 May 1991	nan
104	NIC	Nicaragua	129.494	4.272.352	NI	NIC	americas	g77	lower_middle_income	7.100	1.661,85	republic	nan
105	NL	Netherlands	37.330	15.568.034	NL	NLD	europe	oecd	high_income	301.900	19.392,30	constitutional monarchy	34,50
106	NOK	North Korea	120.540	23.904.124	KP	PRK	asia	g77	low_income	21.500	899,43	Communist state	nan
107	NZ	New Zealand	268.680	3.547.983	NZ	NZL	asia	oecd	high_income	62.300	17.559,27	parliamentary democracy	21,60
108	OM	Oman	212.460	2.186.548	OM	OMN	asia	g77	high_income	19.100	8.735,23	monarchy	nan
109	P	Portugal	92.080	9.865.114	PT	PRT	europe	oecd	high_income	116.200	11.778,88	republic	23,70
110	PA	Panama	78.200	2.655.094	PA	PAN	americas	g77	high_income	13.600	5.122,23	constitutional republic	nan
111	PAL	Palau	458	16.952	PW	PLW	asia	others	high_income	82	4.825,39	constitutional government in free association with	nan
112	PE	Peru	1.285.220	24.523.408	PE	PER	americas	g77	upper_middle_income	87.000	3.547,63	republic	nan
113	PK	Pakistan	803.940	129.275.660	PK	PAK	asia	g77	lower_middle_income	274.200	2.121,05	republic	10,10
114	PL	Poland	312.683	38.642.565	PL	POL	europe	oecd	high_income	226.700	5.866,59	democratic state	24,80
115	PNG	Papua New Guinea	461.690	4.394.537	PG	PNG	asia	g77	lower_middle_income	10.200	2.321,06	parliamentary democracy	nan
116	PY	Paraguay	406.750	5.504.146	PY	PRY	americas	g77	upper_middle_income	17.000	3.088,58	republic	nan
117	Q	Qatar	11.000	547.761	QA	QAT	asia	g77	high_income	10.700	19.534,07	traditional monarchy	19,80
118	R	Russia	17.075.200	148.178.487	RU	RUS	europe	others	upper_middle_income	796.000	5.371,90	federation	23,70
119	RA	Argentina	2.766.890	34.672.997	AR	ARG	americas	g77	upper_middle_income	278.500	8.032,19	republic	17,50
120	RB	Botswana	600.370	1.477.630	BW	BWA	africa	g77	upper_middle_income	4.500	3.045,42	parliamentary republic	nan
121	RC	Taiwan	36.179	21.465.881	None	None	None	None	None	290.500	13.533,10	multiparty democratic regime	25,40
122	RCA	Central African Republic	622.980	3.274.426	CF	CAF	africa	g77	low_income	2.500	763,49	republic	nan
123	RCH	Chile	756.950	14.333.258	CL	CHL	americas	g77	high_income	113.200	7.897,72	republic	20,20

	code	name	area	population	ISO2	ISO3	world_4region	club	income_groups	GDP	GDPPC	Government	overall
124	RG	Guinea	245.860	7.411.981	GN	GIN	africa	g77	low_income	6.500	876,96	republic	nan
125	RH	Haiti	27.750	6.731.539	HT	HTI	americas	g77	low_income	6.500	965,60	republic	nan
126	RI	Indonesia	1.919.440	206.611.600	ID	IDN	asia	g77	lower_middle_income	710.900	3.440,76	republic	18,20
127	RIM	Mauritania	1.030.700	2.336.048	MR	MRT	africa	g77	lower_middle_income	2.800	1.198,61	republic	nan
128	RL	Lebanon	10.400	3.776.317	LB	LBN	asia	g77	upper_middle_income	18.300	4.845,99	republic	nan
129	RM	Madagascar	587.041	13.670.507	MG	MDG	africa	g77	low_income	11.400	833,91	republic	nan
130	RMM	Mali	1.240.000	9.653.261	ML	MLI	africa	g77	low_income	5.400	559,40	republic	nan
131	RN	Niger	1.267.000	9.113.001	NE	NER	africa	g77	low_income	5.500	603,53	republic	nan
132	RO	Romania	237.500	21.657.162	RO	ROU	europe	others	upper_middle_income	105.700	4.880,60	republic	nan
133	ROK	South Korea	98.480	45.482.291	KR	KOR	asia	oecd	high_income	590.700	12.987,47	republic	40,30
134	ROU	Uruguay	176.220	3.238.952	UY	URY	americas	g77	high_income	24.400	7.533,30	republic	16,30
135	RP	Philippines	299.764	74.480.848	PH	PHL	asia	g77	lower_middle_income	179.700	2.412,70	republic	nan
136	RSA	South Africa	1.219.910	41.743.459	ZA	ZAF	africa	g77	upper_middle_income	215.000	5.150,51	republic	14,10
137	RSM	San Marino	60	24.521	SM	SMR	europe	others	high_income	380	15.496,92	republic	nan
138	RT	Togo	56.790	4.570.530	TG	TGO	africa	g77	low_income	4.100	897,05	republic under transition to multiparty democratic	nan
139	RWA	Rwanda	26.340	6.853.359	RW	RWA	africa	g77	low_income	3.800	554,47	republic	nan
140	S	Sweden	449.964	8.900.954	SE	SWE	europe	oecd	high_income	177.300	19.919,21	constitutional monarchy	30,30
141	SA	Saudi Arabia	1.960.580	19.409.058	SA	SAU	asia	g77	high_income	189.300	9.753,18	monarchy	23,30
142	SF	Finland	337.030	5.105.230	FI	FIN	europe	oecd	high_income	92.400	18.099,09	republic	34,90
143	SGP	Singapore	633	3.396.924	SG	SGP	asia	g77	high_income	66.100	19.458,78	republic within Commonwealth	49,70
144	SK	Slovakia	48.845	5.374.362	SK	SVK	europe	oecd	high_income	39.000	7.256,68	parliamentary democracy	17,10
145	SLB	Solomon Islands	28.450	412.902	SB	SLB	asia	g77	lower_middle_income	1.000	2.421,88	parliamentary democracy	nan
146	SLO	Slovenia	20.256	1.951.443	SI	SVN	europe	others	high_income	22.600	11.581,17	emerging democracy	21,50
147	SME	Suriname	163.270	436.418	SR	SUR	americas	g77	upper_middle_income	1.300	2.978,80	republic	nan
148	SN	Senegal	196.190	9.092.749	SN	SEN	africa	g77	lower_middle_income	14.500	1.594,68	republic under multiparty democratic rule	nan
149	SP	Somalia	637.660	9.639.151	SO	SOM	africa	g77	low_income	3.600	373,48	none	nan
150	SRB	Serbia	77.474	7.379.339	RS	SRB	europe	others	upper_middle_income	52.180	7.071,09	parliamentary democracy	nan
151	STP	Sao Tome and Principe	1.001	144.128	ST	STP	africa	g77	lower_middle_income	138	957,48	republic	nan
152	SUD	Sudan	2.505.810	31.547.543	SD	SDN	africa	g77	lower_middle_income	25.000	792,45	transitional previously ruling military junta	nan
153	SY	Seychelles	455	77.575	SC	SYC	africa	g77	high_income	430	5.543,02	republic	nan
154	SYR	Syria	185.180	15.608.648	SY	SYR	asia	g77	low_income	91.200	5.842,91	republic under military regime since March 1963	nan

code	name	area	population	ISO2	ISO3	world_4region	club	income_groups	GDP	GDPPC	Government	overall
155	TAD	Tajikistan	143.100	5.916.373	TJ	TJK	asia	others	low_income	6.400	1.081,74	republic nan
156	TCH	Chad	1.284.000	6.976.845	TD	TCD	africa	g77	low_income	3.300	472,99	republic nan
157	THA	Thailand	514.000	58.851.357	TH	THA	asia	g77	upper_middle_income	416.700	7.080,55	constitutional monarchy nan
158	TJ	China	9.596.960	1.210.004.956	CN	CHN	asia	g77	upper_middle_income	3.500.000	2.892,55	Communist state 61,50
159	TL	Timor-Leste	15.007	1.131.612	TL	TLS	asia	g77	lower_middle_income	489	432,13	parliamentary democracy nan
160	TM	Turkmenistan	488.100	4.149.283	TM	TKM	asia	g77	upper_middle_income	11.500	2.771,56	republic nan
161	TN	Tunisia	163.610	9.019.687	TN	TUN	africa	g77	lower_middle_income	37.100	4.113,22	republic 13,70
162	TO	Tonga	748	106.466	TO	TON	asia	g77	upper_middle_income	228	2.141,53	hereditary constitutional monarchy nan
163	TR	Turkey	780.580	62.484.478	TR	TUR	europe	oecd	upper_middle_income	345.700	5.532,57	republican parliamentary democracy 20,60
164	TT	Trinidad and Tobago	5.130	1.272.385	TT	TTO	americas	g77	high_income	16.200	12.732,00	parliamentary democracy nan
165	TUV	Tuvalu	26	10.146	TV	TUV	asia	others	upper_middle_income	8	768,78	democracy nan
166	UA	Ukraine	603.700	50.864.009	UA	UKR	europe	others	lower_middle_income	174.600	3.432,68	republic nan
167	UAE	United Arab Emirates	75.581	3.057.337	AE	ARE	asia	g77	high_income	70.100	22.928,45	federation with specified powers delegated to the 23,90
168	USA	United States	9.372.610	266.476.278	US	USA	americas	oecd	high_income	7.247.700	27.198,29	federal republic 100,00
169	UZB	Uzbekistan	447.400	23.418.381	UZ	UZB	asia	others	lower_middle_income	54.700	2.335,77	republic nan
170	V	Holy See	0	840	VA	VAT	europe	others	None	nan	nan	monarchical sacerdotal state nan
171	VN	Vietnam	329.560	73.976.973	VN	VNM	asia	g77	lower_middle_income	97.000	1.311,22	Communist state 18,00
172	VU	Vanuatu	14.760	177.504	VU	VUT	asia	g77	lower_middle_income	210	1.183,07	republic nan
173	WAG	Gambia	11.300	1.204.984	GM	GMB	africa	g77	low_income	1.100	912,88	republic under multiparty democratic rule nan
174	WAL	Sierra Leone	71.740	4.793.121	SL	SLE	africa	g77	low_income	4.400	917,98	constitutional democracy nan
175	WAN	Nigeria	923.770	103.912.489	NG	NGA	africa	g77	lower_middle_income	135.900	1.307,83	military government 9,30
176	WD	Dominica	746	82.926	DM	DMA	americas	g77	upper_middle_income	200	2.411,79	parliamentary democracy nan
177	WG	Grenada	344	94.961	GD	GRD	americas	g77	upper_middle_income	284	2.990,70	parliamentary democracy nan
178	WS	Samoa	2.860	214.384	WS	WSM	asia	g77	upper_middle_income	415	1.935,78	constitutional monarchy under native chief nan
179	YE	Yemen	527.970	13.483.178	YE	YEM	asia	g77	low_income	37.100	2.751,58	republic nan
180	YV	Venezuela	912.050	21.983.188	VE	VEN	americas	g77	upper_middle_income	195.500	8.893,16	republic nan
181	Z	Zambia	752.610	9.159.072	ZM	ZMB	africa	g77	lower_middle_income	8.900	971,71	republic nan
182	ZW	Zimbabwe	390.580	11.271.314	ZW	ZWE	africa	g77	lower_middle_income	18.100	1.605,85	parliamentary democracy nan

**Query 7 Results:** The data looks correct, and we can now execute more queries on this view. Perhaps now is a good time to turn our attention to the data gathered from Tortoise Media.

**Query 8:** As stated earlier, we use the Tortoise Media data as a benchmark against which data from other sources can be analyzed. To verify the worth of this data, you might want to [have a look at their web page](#) where there are some beautiful visualizations of the data they obtained. The main data we use from Tortoise Media is in the ranking table. We need to import this data –

including the different colors that are displayed – by web scraping (which was done manually) into an SQL table.

```
In [17]: # Modified function to apply background color only to specified columns (based on main grouping in ai_tortoise_indicators)
def apply_background_color(s, color_map):
    return ['background-color: ' + color_map.get(col, '') for col in s.index]
# Map of columns to colors
color_map = {
    "Overall": "yellow",
    "Talent": "lime",
    "Infrastructure": "lime",
    "Operating environment": "lime",
    "Research": "green",
    "Development": "green",
    "Government strategy": "#10efff",
    "Commercial": "#10efff"
}

query = f"SELECT * FROM ai_tortoise_indicators ORDER BY overall DESC"

try:
    df = pd.read_sql_query(query, engine)
    # Apply formatting and coloring using the modified function
    formatted_df = df.style.apply(apply_background_color, color_map=color_map, axis=1)\n        .format(precision=1, thousands=",", decimal=",")
    # Display the formatted DataFrame
    display(formatted_df)
except Exception as e:
    print(e)
```

	Country	Overall	Talent	Infrastructure	Operating environment	Research	Development	Government strategy	Commercial	Scale	Intensity
0	United States	100,0	100,0	100,0	82,8	100,0	100,0	90,3	100,0	100,0	60,5
1	China	61,5	30,0	92,1	99,7	54,7	80,6	93,5	43,1	67,8	39,3
2	Singapore	49,7	56,9	82,8	85,7	48,8	24,4	81,8	26,2	28,1	100,0
3	United Kingdom	41,8	53,8	61,8	79,5	38,1	19,8	89,2	20,0	34,6	52,8
4	Canada	40,3	46,0	62,1	93,1	34,0	18,9	93,4	18,9	31,9	56,9
5	South Korea	40,3	35,1	74,4	91,4	24,3	60,9	91,9	8,3	31,3	57,3
6	Israel	40,0	45,5	60,5	85,1	24,8	22,2	31,8	40,5	23,5	76,1
7	Germany	39,2	57,0	68,2	90,7	29,3	19,5	93,9	10,3	34,7	46,4
8	Switzerland	37,7	44,5	68,0	81,9	41,3	24,9	9,0	13,3	23,5	71,6
9	Finland	34,9	34,5	73,0	97,7	27,4	13,1	82,7	9,5	24,2	62,0
10	Netherlands	34,5	45,2	65,7	90,3	27,1	15,7	71,8	7,9	25,5	55,5
11	Japan	33,9	38,0	80,8	92,4	18,6	22,2	80,3	6,8	32,2	36,4
12	France	32,8	41,5	68,9	84,2	21,4	8,9	87,3	10,8	29,5	39,7
13	India	31,4	86,2	34,7	91,1	12,0	7,6	56,0	8,9	33,7	20,0
14	Australia	30,9	34,2	54,3	53,8	34,4	11,7	83,3	7,0	23,6	47,5
15	Denmark	30,5	30,6	67,5	100,0	19,4	8,8	75,0	8,3	23,3	50,9
16	Sweden	30,3	33,7	62,2	99,9	22,4	11,4	47,0	8,6	22,5	50,1
17	Luxembourg	29,2	24,0	74,9	90,4	19,4	7,5	66,8	8,8	19,0	53,1
18	Ireland	28,8	31,8	60,8	88,0	13,6	19,2	71,7	8,6	20,1	49,6
19	Austria	27,7	27,6	57,0	94,7	21,3	7,4	63,1	5,8	21,7	44,0
20	Spain	27,7	31,5	65,2	90,2	14,5	4,3	93,4	4,7	24,6	34,5
21	Belgium	26,6	27,4	55,3	84,8	21,9	5,1	57,0	6,6	21,6	41,4
22	Italy	26,5	28,5	56,6	93,7	16,5	3,2	89,8	3,7	24,0	31,6
23	Norway	26,4	27,7	62,1	82,4	16,5	8,1	55,8	7,8	19,5	45,5
24	Estonia	26,0	22,0	57,2	88,3	8,5	2,9	78,8	17,5	17,7	44,9
25	Taiwan	25,4	24,5	71,3	71,4	13,6	16,3	51,4	4,3	20,9	36,7
26	Poland	24,8	31,9	59,0	87,8	10,1	2,2	89,1	2,7	22,0	31,2
27	United Arab Emirates	23,9	16,9	80,8	77,8	9,4	1,7	72,8	5,3	20,0	32,5
28	Russia	23,7	25,1	64,1	80,7	7,9	6,7	91,3	1,7	22,1	24,9
29	Portugal	23,7	20,3	56,4	92,6	11,4	2,2	72,6	6,6	19,4	34,7
30	Saudi Arabia	23,3	14,3	63,3	88,1	8,2	1,3	100,0	6,0	20,6	29,5

	Country	Overall	Talent	Infrastructure	Operating environment	Research	Development	Government strategy	Commercial	Scale	Intensity
31	Hong Kong	22,5	14,4	69,3	80,2	7,0	0,4	15,4	19,2	20,2	32,9
32	Malta	22,4	17,3	56,3	87,7	6,5	12,8	72,6	4,3	15,3	36,7
33	Brazil	22,1	29,3	55,8	75,7	8,5	1,9	71,4	3,3	21,2	22,7
34	Czech Republic	22,1	20,8	53,6	80,4	11,4	2,7	81,2	2,8	18,1	31,9
35	New Zealand	21,6	23,0	58,3	74,8	14,4	4,0	25,3	4,9	16,4	35,7
36	Slovenia	21,5	10,9	59,7	79,9	13,3	2,5	74,8	2,7	16,4	33,3
37	Hungary	20,7	19,0	59,1	83,1	8,1	1,5	59,1	3,4	17,9	28,9
38	Iceland	20,6	16,2	56,1	74,8	15,6	1,7	13,9	8,1	13,8	37,4
39	Turkey	20,6	25,0	45,5	93,6	6,9	0,3	72,5	2,0	19,0	23,3
40	Chile	20,2	15,0	67,5	71,6	3,8	0,6	76,3	5,2	18,6	26,4
41	Qatar	19,8	3,5	61,0	73,1	19,9	0,5	39,2	1,5	15,2	30,9
42	Lithuania	19,7	19,2	56,1	80,2	5,3	2,2	61,0	3,5	15,4	29,1
43	Malaysia	19,6	19,4	65,3	72,2	6,8	0,7	48,1	2,4	17,4	25,4
44	Greece	18,3	26,7	46,5	60,5	13,3	1,8	15,4	2,9	14,0	28,6
45	Indonesia	18,2	28,1	40,9	80,0	3,9	0,3	55,0	2,5	18,0	19,5
46	Vietnam	18,0	21,5	56,2	64,5	3,2	0,7	65,1	1,9	17,6	20,0
47	Colombia	17,8	17,1	52,2	69,4	2,0	0,6	87,7	1,4	16,4	20,6
48	Argentina	17,5	17,5	52,5	80,9	2,6	0,4	56,7	1,4	15,8	20,0
49	Slovakia	17,1	13,6	54,1	84,1	3,4	0,6	43,5	1,9	14,1	22,5
50	Mexico	16,9	20,2	43,8	78,2	4,2	0,4	53,5	1,1	16,1	18,7
51	Egypt	16,9	18,5	42,7	66,6	4,7	0,3	74,8	1,5	16,3	17,9
52	Uruguay	16,3	13,3	60,8	79,7	0,8	1,1	31,5	2,1	13,6	21,0
53	Armenia	14,5	21,6	47,8	71,5	1,6	2,0	0,0	2,3	11,7	21,5
54	South Africa	14,1	11,4	41,4	81,5	3,8	0,3	0,0	5,4	13,7	17,1
55	Tunisia	13,7	17,8	43,8	70,9	3,9	0,3	4,9	0,8	11,4	18,8
56	Morocco	13,6	9,1	52,7	73,1	2,4	0,1	16,4	0,7	12,1	15,1
57	Bahrain	13,5	4,4	56,2	77,2	2,0	0,1	0,0	3,6	11,4	17,8
58	Pakistan	10,1	22,9	9,0	55,7	3,9	0,2	13,4	1,3	10,3	10,6
59	Sri Lanka	10,0	11,8	34,9	56,0	0,2	0,0	4,9	1,5	9,0	11,6
60	Nigeria	9,3	16,6	5,0	64,6	3,5	0,1	13,9	1,3	8,0	10,7
61	Kenya	8,3	10,6	13,5	59,3	1,3	0,1	9,0	2,1	7,6	9,4

**Query 8 Results:** As you can see, there are 62 countries used in this table and the following 3 ranking groups:

- implementation: talent, infrastructure, operating environment (lime)
- innovation: research, development (green)
- investment: government strategy, commercial (blue)

We will use this data to verify whether collected data from other tables supports Tortoise Media's findings.

## 5.2. Economics and Individual Countries

**Query 9:** Importantly, we have already included the score from Tortoise Media's ranking table in our created view. We can now check whether there is a correlation between any of the general parameters – such as GDP, GDP/capita, population – and this overall ranking parameter.

*Note: population and GDP data are based on figures from the countries table. These are somewhat dated, but it remains interesting to see if data from about 10 years ago has relevance in the prognoses of technological development in the future. When more thorough research can be done, the GDP evolution over time could be utilized to develop a more up-to-date comparison. For our current purposes, these figures will suffice.*

```
In [18]: # RANK() OVER (ORDER BY ...) returns the ranking within the dataframe following the "ORDER BY" clause between round brackets
query = """
SELECT Name, Population, GDPPC, GDP,
RANK() OVER ( ORDER BY Population DESC) population_rank,
RANK() OVER ( ORDER BY GDPPC DESC) GDPPC_rank,
RANK() OVER ( ORDER BY GDP DESC) GDP_rank,
RANK() OVER ( ORDER BY overall DESC) overall_rank
FROM {view_name}
ORDER BY GDP DESC
"""

try:
    # Apply formatting and coloring using the modified function
    df = pd.read_sql_query(query, engine)
    formatted_df = df.style\
        .format(precision=0, thousands=",", decimal=",")
    # Display the formatted DataFrame
    display(formatted_df)
except Exception as e:
    print(e)
```

	name	population	GDPPC	GDP	population_rank	GDPPC_rank	GDP_rank	overall_rank
0	United States	266.476.278	27.198	7.247.700	3	2	1	1
1	China	1.210.004.956	2.893	3.500.000	1	101	2	2
2	Japan	125.449.703	21.357	2.679.200	8	11	3	12
3	Germany	83.536.115	17.384	1.452.200	12	26	4	8
4	India	952.107.694	1.480	1.408.700	2	127	5	14
5	France	58.317.450	20.114	1.173.000	20	13	6	13
6	United Kingdom	58.489.975	19.463	1.138.400	19	16	7	4
7	Italy	57.460.274	18.945	1.088.600	21	20	8	23
8	Brazil	162.661.214	6.005	976.800	5	64	9	34
9	Russia	148.178.487	5.372	796.000	6	70	10	29
10	Mexico	95.772.462	7.532	721.400	11	55	11	51
11	Indonesia	206.611.600	3.441	710.900	4	91	12	46
12	Canada	28.820.671	24.080	694.000	35	5	13	5
13	South Korea	45.482.291	12.987	590.700	25	35	14	5
14	Spain	39.181.114	14.420	565.000	27	32	15	20
15	Thailand	58.851.357	7.081	416.700	18	58	16	63
16	Australia	18.260.863	22.200	405.400	50	8	17	15
17	Turkey	62.484.478	5.533	345.700	17	69	18	39
18	Iran	66.094.264	4.895	323.500	15	77	19	63
19	Hong Kong	7.055.071	43.600	307.600	88	1	20	32
20	Netherlands	15.568.034	19.392	301.900	54	18	21	11
21	Taiwan	21.465.881	13.533	290.500	44	34	22	26
22	Argentina	34.672.997	8.032	278.500	30	52	23	49
23	Pakistan	129.275.660	2.121	274.200	7	116	24	59
24	Poland	38.642.565	5.867	226.700	28	66	25	27
25	South Africa	41.743.459	5.151	215.000	26	73	26	55
26	Belgium	10.170.241	19.370	197.000	69	19	27	22
27	Venezuela	21.983.188	8.893	195.500	42	49	28	63
28	Malaysia	19.962.893	9.698	193.600	47	46	29	44
29	Colombia	36.813.161	5.229	192.500	29	72	30	48
30	Saudi Arabia	19.409.058	9.753	189.300	48	43	31	31

	name	population	GDPPC	GDP	population_rank	GDPPC_rank	GDP_rank	overall_rank
31	Philippines	74.480.848	2.413	179.700	13	109	32	63
32	Sweden	8.900.954	19.919	177.300	79	14	33	17
33	Ukraine	50.864.009	3.433	174.600	23	92	34	63
34	Egypt	63.575.107	2.690	171.000	16	105	35	51
35	Switzerland	7.207.060	21.992	158.500	86	9	36	9
36	Austria	8.023.244	18.945	152.000	82	21	37	20
37	Bangladesh	123.062.800	1.174	144.500	9	140	38	63
38	Nigeria	103.912.489	1.308	135.900	10	134	39	61
39	Portugal	9.865.114	11.779	116.200	71	38	40	29
40	Chile	14.333.258	7.898	113.200	55	53	41	41
41	Denmark	5.249.632	21.487	112.800	101	10	42	16
42	Algeria	29.183.032	3.725	108.700	33	86	43	63
43	Norway	4.383.807	24.226	106.200	109	4	44	24
44	Czech Republic	10.321.120	10.290	106.200	68	42	44	34
45	Romania	21.657.162	4.881	105.700	43	78	46	63
46	Greece	10.538.594	9.650	101.700	65	47	47	45
47	Vietnam	73.976.973	1.311	97.000	14	133	48	47
48	Finland	5.105.230	18.099	92.400	103	24	49	10
49	Syria	15.608.648	5.843	91.200	53	67	50	63
50	Morocco	29.779.156	2.935	87.400	32	100	51	57
51	Peru	24.523.408	3.548	87.000	37	88	52	63
52	Israel	5.421.995	14.773	80.100	99	31	53	7
53	Hungary	10.002.541	7.248	72.500	70	57	54	38
54	United Arab Emirates	3.057.337	22.928	70.100	124	7	55	28
55	Singapore	3.396.924	19.459	66.100	120	17	56	3
56	Sri Lanka	18.553.074	3.536	65.600	49	89	57	60
57	New Zealand	3.547.983	17.559	62.300	116	25	58	36
58	Uzbekistan	23.418.381	2.336	54.700	39	111	59	63
59	Ireland	3.566.833	15.308	54.600	115	30	60	19
60	Serbia	7.379.339	7.071	52.180	85	59	61	63
61	Belarus	10.415.973	4.724	49.200	66	81	62	63

	<b>name</b>	<b>population</b>	<b>GDPPC</b>	<b>GDP</b>	<b>population_rank</b>	<b>GDPPC_rank</b>	<b>GDP_rank</b>	<b>overall_rank</b>
<b>62</b>	Myanmar	45.975.625	1.022	47.000	24	146	63	63
<b>63</b>	Ecuador	11.466.291	3.890	44.600	59	85	64	63
<b>64</b>	Bulgaria	8.612.757	5.016	43.200	80	76	65	63
<b>65</b>	Iraq	21.422.292	1.919	41.100	45	121	66	63
<b>66</b>	Slovakia	5.374.362	7.257	39.000	100	56	67	50
<b>67</b>	Tunisia	9.019.687	4.113	37.100	78	83	68	56
<b>68</b>	Yemen	13.483.178	2.752	37.100	58	104	68	63
<b>69</b>	Kenya	28.176.686	1.306	36.800	36	135	70	62
<b>70</b>	Guatemala	11.277.614	3.254	36.700	60	94	71	63
<b>71</b>	Libya	5.445.436	6.042	32.900	98	62	72	63
<b>72</b>	Kuwait	1.950.047	15.794	30.800	135	27	73	63
<b>73</b>	Dominican Republic	8.088.881	3.313	26.800	81	93	74	63
<b>74</b>	Nepal	22.094.033	1.141	25.200	41	143	75	63
<b>75</b>	Ghana	17.698.271	1.418	25.100	52	130	76	63
<b>76</b>	Sudan	31.547.543	792	25.000	31	161	77	63
<b>77</b>	Uruguay	3.238.952	7.533	24.400	123	54	78	53
<b>78</b>	Ethiopia	57.171.662	423	24.200	22	180	79	63
<b>79</b>	Tanzania	29.058.470	795	23.100	34	160	80	63
<b>80</b>	Slovenia	1.951.443	11.581	22.600	134	40	81	37
<b>81</b>	North Korea	23.904.124	899	21.500	38	153	82	63
<b>82</b>	Croatia	5.004.112	4.017	20.100	104	84	83	63
<b>83</b>	Bolivia	7.165.257	2.791	20.000	87	102	84	63
<b>84</b>	Jordan	4.212.152	4.582	19.300	111	82	85	63
<b>85</b>	Oman	2.186.548	8.735	19.100	131	50	86	63
<b>86</b>	Costa Rica	3.463.083	5.313	18.400	118	71	87	63
<b>87</b>	Lebanon	3.776.317	4.846	18.300	113	79	88	63
<b>88</b>	Zimbabwe	11.271.314	1.606	18.100	61	124	89	63
<b>89</b>	Paraguay	5.504.146	3.089	17.000	97	96	90	63
<b>90</b>	Uganda	20.158.176	833	16.800	46	159	91	63
<b>91</b>	Cameroon	14.261.557	1.157	16.500	56	142	92	63
<b>92</b>	Trinidad and Tobago	1.272.385	12.732	16.200	140	36	93	63

	name	population	GDPPC	GDP	population_rank	GDPPC_rank	GDP_rank	overall_rank
93	Latvia	2.468.982	5.954	14.700	129	65	94	63
94	Cuba	10.951.334	1.342	14.700	62	131	94	63
95	Senegal	9.092.749	1.595	14.500	77	125	96	63
96	Panama	2.655.094	5.122	13.600	126	74	97	63
97	Lithuania	3.646.041	3.648	13.300	114	87	98	43
98	Afghanistan	22.664.136	565	12.800	40	175	99	63
99	Estonia	1.459.428	8.428	12.300	139	51	100	25
100	Mozambique	17.877.927	682	12.200	51	169	101	63
101	Turkmenistan	4.149.283	2.772	11.500	112	103	102	63
102	Azerbaijan	7.676.953	1.498	11.500	83	126	102	63
103	El Salvador	5.828.987	1.956	11.400	94	118	104	63
104	Madagascar	13.670.507	834	11.400	57	158	104	63
105	Mauritius	1.140.256	9.559	10.900	144	48	106	63
106	Honduras	5.605.193	1.927	10.800	96	120	107	63
107	Qatar	547.761	19.534	10.700	152	15	108	42
108	Moldova	4.463.847	2.330	10.400	107	112	109	63
109	Papua New Guinea	4.394.537	2.321	10.200	108	113	110	63
110	Luxembourg	415.870	24.046	10.000	157	6	111	18
111	Armenia	3.463.574	2.627	9.100	117	106	112	54
112	Zambia	9.159.072	972	8.900	75	148	113	63
113	Jamaica	2.595.275	3.160	8.200	127	95	114	63
114	Cyprus	744.609	10.475	7.800	147	41	115	63
115	Benin	5.709.529	1.331	7.600	95	132	116	63
116	Angola	10.342.899	715	7.400	67	166	117	63
117	Burkina Faso	10.623.323	697	7.400	64	168	117	63
118	Bahrain	590.042	12.372	7.300	150	37	119	58
119	Nicaragua	4.272.352	1.662	7.100	110	122	120	63
120	Cambodia	10.861.218	644	7.000	63	172	121	63
121	Malawi	9.452.844	730	6.900	74	165	122	63
122	Haiti	6.731.539	966	6.500	91	149	123	63
123	Guinea	7.411.981	877	6.500	84	155	123	63

	<b>name</b>	<b>population</b>	<b>GDPPC</b>	<b>GDP</b>	<b>population_rank</b>	<b>GDPPC_rank</b>	<b>GDP_rank</b>	<b>overall_rank</b>
<b>124</b>	Tajikistan	5.916.373	1.082	6.400	93	145	125	63
<b>125</b>	Georgia	5.219.810	1.188	6.200	102	138	126	63
<b>126</b>	Gabon	1.172.798	5.116	6.000	142	75	127	63
<b>127</b>	Namibia	1.677.243	3.458	5.800	137	90	128	63
<b>128</b>	Niger	9.113.001	604	5.500	76	173	129	63
<b>129</b>	Mali	9.653.261	559	5.400	72	176	130	63
<b>130</b>	Iceland	270.292	18.499	5.000	162	23	131	39
<b>131</b>	Mongolia	2.496.617	1.963	4.900	128	117	132	63
<b>132</b>	Bahamas	259.367	18.507	4.800	163	22	133	63
<b>133</b>	Fiji	782.381	6.007	4.700	146	63	134	63
<b>134</b>	Brunei	299.939	15.336	4.600	160	29	135	63
<b>135</b>	Montenegro	672.180	6.717	4.515	149	60	136	63
<b>136</b>	Botswana	1.477.630	3.045	4.500	138	97	137	63
<b>137</b>	Malta	375.576	11.715	4.400	159	39	138	33
<b>138</b>	Sierra Leone	4.793.121	918	4.400	105	151	138	63
<b>139</b>	Albania	3.249.136	1.262	4.100	122	136	140	63
<b>140</b>	Togo	4.570.530	897	4.100	106	154	140	63
<b>141</b>	Burundi	5.943.057	673	4.000	92	170	142	63
<b>142</b>	Rwanda	6.853.359	554	3.800	90	177	143	63
<b>143</b>	Somalia	9.639.151	373	3.600	73	182	144	63
<b>144</b>	Chad	6.976.845	473	3.300	89	178	145	63
<b>145</b>	Lesotho	1.970.781	1.421	2.800	133	129	146	63
<b>146</b>	Mauritania	2.336.048	1.199	2.800	130	137	146	63
<b>147</b>	Barbados	257.030	9.726	2.500	164	45	148	63
<b>148</b>	Central African Republic	3.274.426	763	2.500	121	163	148	63
<b>149</b>	Liberia	2.109.789	1.090	2.300	132	144	150	63
<b>150</b>	Eritrea	3.427.883	583	2.000	119	174	151	63
<b>151</b>	Guyana	712.091	2.247	1.600	148	114	152	63
<b>152</b>	Suriname	436.418	2.979	1.300	154	99	153	63
<b>153</b>	Bhutan	1.822.625	713	1.300	136	167	153	63
<b>154</b>	Gambia	1.204.984	913	1.100	141	152	155	63

	name	population	GDPPC	GDP	population_rank	GDPPC_rank	GDP_rank	overall_rank
155	Andorra	72.766	13.743	1.000	174	33	156	63
156	Solomon Islands	412.902	2.422	1.000	158	108	156	63
157	Guinea-Bissau	1.151.330	869	1.000	143	156	156	63
158	Bosnia and Herzegovina	2.656.240	376	1.000	125	181	156	63
159	Monaco	31.719	24.843	788	177	3	160	63
160	Liechtenstein	31.122	20.243	630	178	12	161	63
161	Belize	219.296	2.622	575	165	107	162	63
162	Djibouti	427.642	1.169	500	156	141	163	63
163	Timor-Leste	1.131.612	432	489	145	179	164	63
164	Cape Verde	449.066	980	440	153	147	165	63
165	Seychelles	77.575	5.543	430	173	68	166	63
166	Antigua and Barbuda	65.647	6.474	425	175	61	167	63
167	Samoa	214.384	1.936	415	166	119	168	63
168	Maldives	270.758	1.440	390	161	128	169	63
169	San Marino	24.521	15.497	380	179	28	170	63
170	Comoros	569.237	650	370	151	171	171	63
171	Equatorial Guinea	431.282	754	325	155	164	172	63
172	Grenada	94.961	2.991	284	170	98	173	63
173	Tonga	106.466	2.142	228	169	115	174	63
174	Vanuatu	177.504	1.183	210	167	139	175	63
175	Dominica	82.926	2.412	200	171	110	176	63
176	Sao Tome and Principe	144.128	957	138	168	150	177	63
177	Nauru	10.273	9.734	100	181	44	178	63
178	Marshall Islands	58.363	1.611	94	176	123	179	63
179	Palau	16.952	4.825	82	180	80	180	63
180	Kiribati	80.919	840	68	172	157	181	63
181	Tuvalu	10.146	769	8	182	162	182	63
182	Holy See	840	nan	nan	183	183	183	63

**Query 9 Results:** As the table shows, the general parameters from our collected data do not directly corroborate the overall rankings defined by Tortoise Media. For instance, Japan has a GDP\_rank of 3, but an overall ai rank of 12. This is interesting and worth exploring further through more queries. But first, let's check the correlation results.

Note: an overall ai rank of 63 indicates that the country has no ranking in ai\_tortoise\_indicators table.

```
In [19]: print("GDP/overall AI rank:", str(round(df['overall_rank'].corr(df['GDP_rank']),2)))
print("GDP per capita/overall AI rank:", str(round(df['overall_rank'].corr(df['GDPPC_rank']),2)))
print("Population/overall AI rank:", str(round(df['overall_rank'].corr(df['population_rank']),2)))

GDP/overall AI rank: 0.62
GDP per capita/overall AI rank: 0.63
Population/overall AI rank: 0.33
```

**Query 10:** There is a moderate but significant correlation between the economic parameters and the AI score. Population means less in this context. To really investigate this correlation further, we might take away all countries that have no overall ranking score.

```
In [20]: query = f"""
SELECT Name, Population, GDPPC, GDP,
RANK() OVER ( ORDER BY Population DESC) population_rank,
RANK() OVER ( ORDER BY GDPPC DESC) GDPPC_rank,
RANK() OVER (ORDER BY GDP DESC) GDP_rank,
RANK() OVER ( ORDER BY overall DESC) overall_rank
FROM {view_name}
WHERE overall IS NOT NULL
ORDER BY GDP DESC
"""

try:
    # Apply formatting and coloring using the modified function
    df = pd.read_sql_query(query, engine)
    formatted_df = df.style\
        .format(precision=0, thousands=",", decimal=",")
    # Display the formatted DataFrame
    display(formatted_df)
except Exception as e:
    print(e)
```

	name	population	GDPPC	GDP	population_rank	GDPPC_rank	GDP_rank	overall_rank
0	United States	266.476.278	27.198	7.247.700	3	2	1	1
1	China	1.210.004.956	2.893	3.500.000	1	55	2	2
2	Japan	125.449.703	21.357	2.679.200	8	10	3	12
3	Germany	83.536.115	17.384	1.452.200	11	23	4	8
4	India	952.107.694	1.480	1.408.700	2	59	5	14
5	France	58.317.450	20.114	1.173.000	16	11	6	13
6	United Kingdom	58.489.975	19.463	1.138.400	15	14	7	4
7	Italy	57.460.274	18.945	1.088.600	17	18	8	23
8	Brazil	162.661.214	6.005	976.800	5	44	9	34
9	Russia	148.178.487	5.372	796.000	6	47	10	29
10	Mexico	95.772.462	7.532	721.400	10	41	11	51
11	Indonesia	206.611.600	3.441	710.900	4	53	12	46
12	Canada	28.820.671	24.080	694.000	25	4	13	5
13	South Korea	45.482.291	12.987	590.700	18	28	14	5
14	Spain	39.181.114	14.420	565.000	20	26	15	20
15	Australia	18.260.863	22.200	405.400	31	7	16	15
16	Turkey	62.484.478	5.533	345.700	14	46	17	39
17	Hong Kong	7.055.071	43.600	307.600	43	1	18	32
18	Netherlands	15.568.034	19.392	301.900	32	16	19	11
19	Taiwan	21.465.881	13.533	290.500	27	27	20	26
20	Argentina	34.672.997	8.032	278.500	23	38	21	49
21	Pakistan	129.275.660	2.121	274.200	7	58	22	59
22	Poland	38.642.565	5.867	226.700	21	45	23	27
23	South Africa	41.743.459	5.151	215.000	19	49	24	55
24	Belgium	10.170.241	19.370	197.000	36	17	25	22
25	Malaysia	19.962.893	9.698	193.600	28	35	26	44
26	Colombia	36.813.161	5.229	192.500	22	48	27	48
27	Saudi Arabia	19.409.058	9.753	189.300	29	34	28	31
28	Sweden	8.900.954	19.919	177.300	40	12	29	17
29	Egypt	63.575.107	2.690	171.000	13	56	30	51
30	Switzerland	7.207.060	21.992	158.500	42	8	31	9

	<b>name</b>	<b>population</b>	<b>GDPPC</b>	<b>GDP</b>	<b>population_rank</b>	<b>GDPPC_rank</b>	<b>GDP_rank</b>	<b>overall_rank</b>
<b>31</b>	Austria	8.023.244	18.945	152.000	41	19	32	20
<b>32</b>	Nigeria	103.912.489	1.308	135.900	9	61	33	61
<b>33</b>	Portugal	9.865.114	11.779	116.200	38	30	34	29
<b>34</b>	Chile	14.333.258	7.898	113.200	33	39	35	41
<b>35</b>	Denmark	5.249.632	21.487	112.800	46	9	36	16
<b>36</b>	Czech Republic	10.321.120	10.290	106.200	35	33	37	34
<b>37</b>	Norway	4.383.807	24.226	106.200	48	3	37	24
<b>38</b>	Greece	10.538.594	9.650	101.700	34	36	39	45
<b>39</b>	Vietnam	73.976.973	1.311	97.000	12	60	40	47
<b>40</b>	Finland	5.105.230	18.099	92.400	47	21	41	10
<b>41</b>	Morocco	29.779.156	2.935	87.400	24	54	42	57
<b>42</b>	Israel	5.421.995	14.773	80.100	44	25	43	7
<b>43</b>	Hungary	10.002.541	7.248	72.500	37	43	44	38
<b>44</b>	United Arab Emirates	3.057.337	22.928	70.100	55	6	45	28
<b>45</b>	Singapore	3.396.924	19.459	66.100	53	15	46	3
<b>46</b>	Sri Lanka	18.553.074	3.536	65.600	30	52	47	60
<b>47</b>	New Zealand	3.547.983	17.559	62.300	51	22	48	36
<b>48</b>	Ireland	3.566.833	15.308	54.600	50	24	49	19
<b>49</b>	Slovakia	5.374.362	7.257	39.000	45	42	50	50
<b>50</b>	Tunisia	9.019.687	4.113	37.100	39	50	51	56
<b>51</b>	Kenya	28.176.686	1.306	36.800	26	62	52	62
<b>52</b>	Uruguay	3.238.952	7.533	24.400	54	40	53	53
<b>53</b>	Slovenia	1.951.443	11.581	22.600	56	32	54	37
<b>54</b>	Lithuania	3.646.041	3.648	13.300	49	51	55	43
<b>55</b>	Estonia	1.459.428	8.428	12.300	57	37	56	25
<b>56</b>	Qatar	547.761	19.534	10.700	59	13	57	42
<b>57</b>	Luxembourg	415.870	24.046	10.000	60	5	58	18
<b>58</b>	Armenia	3.463.574	2.627	9.100	52	57	59	54
<b>59</b>	Bahrain	590.042	12.372	7.300	58	29	60	58
<b>60</b>	Iceland	270.292	18.499	5.000	62	20	61	39
<b>61</b>	Malta	375.576	11.715	4.400	61	31	62	33

```
In [21]: print("GDP/overall AI rank:", str(round(df['overall_rank'].corr(df['GDP_rank']),2)))
print("GDP per capita/overall AI rank:", str(round(df['overall_rank'].corr(df['GDPPC_rank']),2)))
print("Population/overall AI rank:", str(round(df['overall_rank'].corr(df['population_rank']),2)))
```

GDP/overall AI rank: 0.44  
 GDP per capita/overall AI rank: 0.67  
 Population/overall AI rank: 0.05

**Query 10 Results:** This query and the correlation calculations make it crystal clear that economic parameters are the most prominent factor in the AI scoring statistics.

**Query 11:** To continue our investigation of the relationship between Tortoise Media and the general country statistics we have gathered, we need to create a new view that makes certain calculations and visualizations easier to understand.

```
In [22]: view_name_economic = 'main_economic_rankings'

# This creates a ranking view
query = f"""CREATE VIEW {view_name_economic} AS
SELECT Name, Population, GDPPC, GDP,
RANK() OVER ( ORDER BY Population DESC) population_rank,
RANK() OVER ( ORDER BY GDPPC DESC) GDPPC_rank,
ROW_NUMBER() OVER (ORDER BY GDP DESC) GDP_rank,
RANK() OVER ( ORDER BY overall DESC) overall_rank
from {view_name} ORDER BY GDP DESC;
"""

try:
    with engine.connect() as conn:
        drop_view_query = f"DROP VIEW IF EXISTS `'{view_name_economic}`" # Using backticks in case of special characters
        conn.execute(text(drop_view_query))
        conn.execute(text(query))
except Exception as e:
    print(e)
```

**Query 11 Results:** Similarly to Query 6, the output is error-free, confirming that our code was inputted correctly.

**Query 12:** With this new view, we now want to see if there are any countries that are not ranked by Tortoise Media (i.e., countries that would be considered 63 on the ranking table) but have a relatively high GDP and GDP/capita. We will limit ourselves to those countries whose GDP and GDP/capita fall within the top 50.

```
In [23]: query = f"SELECT * FROM {view_name_economic} WHERE overall_rank = 63 AND GDP_rank < 51 AND GDPPC_rank < 51 ORDER BY GDP DESC"

try:
    # Apply formatting using the modified function
    df = pd.read_sql_query(query, engine)
    formatted_df = df.style\
        .format(precision=0, thousands=",", decimal=",")
    # Display the formatted DataFrame
    display(formatted_df)
except Exception as e:
    print(e)
```

	Name	Population	GDPPC	GDP	population_rank	GDPPC_rank	GDP_rank	overall_rank
0	Venezuela	21.983.188	8.893	195.500	42	49	28	63

**Query 12 Results:** The only country that meets these criteria is Venezuela. It could be interesting to further investigate why this country is not ranked in the benchmark table; however, for this project, we will simply note it and move on.

**Query 13:** Perhaps it is just as relevant to look for countries that have been ranked (i.e., countries that are ranked 62 or lower on Tortoise Media's ranking table) and have GDP and GDP/capita that falls outside the top 50.

```
In [24]: query = f"SELECT * FROM {view_name_economic} WHERE overall_rank < 63 AND GDP_rank > 50 AND GDPPC_rank > 50 ORDER BY overall_rank"
try:
    # Apply formatting using the modified function
    df = pd.read_sql_query(query, engine)
    formatted_df = df.style\
        .format(precision=0, thousands=",", decimal=",")
    # Display the formatted DataFrame
    display(formatted_df)
except Exception as e:
    print(e)
```

	Name	Population	GDPPC	GDP	population_rank	GDPPC_rank	GDP_rank	overall_rank
0	Estonia	1.459.428	8.428	12.300	139	51	100	25
1	Hungary	10.002.541	7.248	72.500	70	57	54	38
2	Lithuania	3.646.041	3.648	13.300	114	87	98	43
3	Slovakia	5.374.362	7.257	39.000	100	56	67	50
4	Uruguay	3.238.952	7.533	24.400	123	54	78	53
5	Armenia	3.463.574	2.627	9.100	117	106	112	54
6	Tunisia	9.019.687	4.113	37.100	78	83	68	56
7	Morocco	29.779.156	2.935	87.400	32	100	51	57
8	Sri Lanka	18.553.074	3.536	65.600	49	89	57	60
9	Kenya	28.176.686	1.306	36.800	36	135	70	62

**Query 13 Results:** Again, the query reveals intriguing information. We see that the first seven countries are all quite small in terms of population and have relatively small economies. Further research into these countries could reveal more information as to why they score well in our overall ranking table.

**Query 14:** While we currently are unable to dig deeper into the countries from the above query, one thing we can do is look more closely at their figures in the benchmark table. To do this in the most legible way, we have grouped the columns by Implementation, Innovation and Investment criteria and averaged the respective columns that fall under each.

```
In [25]: query = """
SELECT Country, Overall,
(Talent+Infrastructure+'Operating environment')/3 AS Implementation,
(Research+Development)/2 AS Innovation,
```

```
('Government strategy'+Commercial)/2 AS Investment
FROM ai_tortoise_indicators
WHERE Country IN
(SELECT Name
FROM {view_name_economic}
WHERE overall_rank < 63 AND GDP_rank > 50 AND GDPPC_rank > 50)
ORDER BY overall DESC;
"""

# Map of columns to colors
color_map2 = {
    "Overall": "yellow",
    "Implementation": "lime",
    "Innovation": "green",
    "Investment": "#10efff"
}

try:
    # Apply formatting using the modified function
    df = pd.read_sql_query(query, engine)
    formatted_df = df.style.apply(apply_background_color, color_map=color_map2, axis=1) \
        .format(precision=1, thousands=",", decimal=",")

    # Display the formatted DataFrame
    display(formatted_df)
except Exception as e:
    print(e)
```

	Country	Overall	Implementation	Innovation	Investment
0	Estonia	26,0	26,4	5,7	8,8
1	Hungary	20,7	26,0	4,8	1,7
2	Lithuania	19,7	25,1	3,8	1,8
3	Slovakia	17,1	22,6	2,0	0,9
4	Uruguay	16,3	24,7	1,0	1,1
5	Armenia	14,5	23,1	1,8	1,1
6	Tunisia	13,7	20,5	2,1	0,4
7	Morocco	13,6	20,6	1,2	0,3
8	Sri Lanka	10,0	15,6	0,1	0,8
9	Kenya	8,3	8,0	0,7	1,1

**Query 14 Results:** Estonia is clearly a special case here. Although it is one of the smaller countries (it has a population of only 1,459,428 inhabitants), it still attracts substantially more public and private capital than its direct neighbors in this table.

**Query 15:** Since Estonia stands out, we want to now see how its Implementation, Innovation and Investment data compares to countries with higher GDPs/overall rankings. To do this, we create a table of countries whose overall rankings are up to 20 places higher than Estonia.

In [26]:

```
query = """
SELECT t1.Country, t1.Overall,
```

```
(Talent+Infrastructure+'Operating environment')/3 AS Implementation,  
(Research+Development)/2 AS Innovation,  
('Government strategy'+Commercial)/2 AS Investment,  
t2.Population, t2.GDP, t2.GDPPC  
FROM ai_tortoise_indicators AS t1  
LEFT JOIN {view_name} AS t2 ON t1.Country = t2.Name  
WHERE Country IN  
(SELECT Name FROM {view_name_economic} WHERE overall_rank BETWEEN 5 AND 25)  
ORDER BY overall DESC;  
"""  
  
# Map of columns to colors  
color_map2 = {  
    "Overall": "yellow",  
    "Implementation": "lime",  
    "Innovation": "green",  
    "Investment": "#10efff"  
}  
  
try:  
    # Apply formatting using the modified function  
    df = pd.read_sql_query(query, engine)  
    formatted_df = df.style.apply(apply_background_color, color_map=color_map2, axis=1)\br/>        .format(precision=1, thousands=",", decimal=".")  
    # Display the formatted DataFrame  
    display(formatted_df)  
except Exception as e:  
    print(e)
```

Country	Overall	Implementation	Innovation	Investment	population	GDP	GDPPC
0 Canada	40,3	36,0	26,4	9,4	28.820.671	694.000,0	24.079,9
1 South Korea	40,3	36,5	42,6	4,2	45.482.291	590.700,0	12.987,5
2 Israel	40,0	35,3	23,5	20,2	5.421.995	80.100,0	14.773,2
3 Germany	39,2	41,7	24,4	5,2	83.536.115	1.452.200,0	17.384,1
4 Switzerland	37,7	37,5	33,1	6,7	7.207.060	158.500,0	21.992,3
5 Finland	34,9	35,8	20,2	4,8	5.105.230	92.400,0	18.099,1
6 Netherlands	34,5	37,0	21,4	4,0	15.568.034	301.900,0	19.392,3
7 Japan	33,9	39,6	20,4	3,4	125.449.703	2.679.200,0	21.356,8
8 France	32,8	36,8	15,1	5,4	58.317.450	1.173.000,0	20.114,0
9 India	31,4	40,3	9,8	4,5	952.107.694	1.408.700,0	1.479,6
10 Australia	30,9	29,5	23,0	3,5	18.260.863	405.400,0	22.200,5
11 Denmark	30,5	32,7	14,1	4,2	5.249.632	112.800,0	21.487,2
12 Sweden	30,3	32,0	16,9	4,3	8.900.954	177.300,0	19.919,2
13 Luxembourg	29,2	33,0	13,4	4,4	415.870	10.000,0	24.046,0
14 Ireland	28,8	30,9	16,4	4,3	3.566.833	54.600,0	15.307,7
15 Spain	27,7	32,2	9,4	2,4	39.181.114	565.000,0	14.420,2
16 Austria	27,7	28,2	14,4	2,9	8.023.244	152.000,0	18.945,0
17 Belgium	26,6	27,6	13,5	3,3	10.170.241	197.000,0	19.370,2
18 Italy	26,5	28,4	9,8	1,9	57.460.274	1.088.600,0	18.945,3
19 Norway	26,4	29,9	12,3	3,9	4.383.807	106.200,0	24.225,5
20 Estonia	26,0	26,4	5,7	8,8	1.459.428	12.300,0	8.428,0

**Query 15 Results:** Estonia's investment still remains impressive, even more so if GDP and population figures are taken into account. Only Israel (ranked 7th overall) and Canada (ranked 5th overall) score better on the investment front.

### 5.3. Geography, Continents, and World Wealth

**Query 16:** Let's now cast a somewhat wider net by grouping countries in our benchmark table according to geographic properties (continent), social markers (income groups), organisational groups (club) or government structures.

```
In [27]: query = f"""
SELECT world_4region AS continent, AVG(overall) AS ai_performance FROM {view_name}
WHERE overall IS NOT NULL
GROUP BY continent
ORDER BY ai_performance DESC;
"""
```

```

try:
    # Execute the query and store the result in a DataFrame
    df = pd.read_sql_query(query, engine)

    # Apply formatting for thousands and decimal points
    formatted_df = df.style \
        .format({
            'area': '{:.0f}',
            'population': '{:.0f}',
            'GDP': '{:.0f}',
            'GDPPC': '{:.2f}',
            'overall': '{:.2f}'
        }, thousands=",", decimal=",")

    # Display the formatted DataFrame
    display(formatted_df)
except Exception as e:
    print(e)

```

	continent	ai_performance
0	americas	31,387500
1	asia	27,394118
2	europe	26,562069
3	None	23,950000
4	africa	12,650000

**Query 16 Results:** What we see is that average performance is reasonably distributed over all continents except Africa. The "None" in continent represents the two Asian subregions of China, namely Taiwan and Hong Kong, that we discussed earlier. If we take this average into account, the Asian average is about on par with the European.

**Query 17 & 18:** Now, let's have a look at the difference in income within the continents. For convenience, we filter first on the higher averages and then on the lower averages.

```

In [28]: query = """
SELECT world_4region AS continent, income_groups, AVG(overall) AS ai_performance
FROM {view_name}
WHERE overall IS NOT NULL
GROUP BY continent, income_groups
HAVING ai_performance > 20
ORDER BY continent, income_groups, ai_performance DESC
"""

try:
    # Execute the query and store the result in a DataFrame
    df = pd.read_sql_query(query, engine)

    # Apply formatting for thousands and decimal points
    formatted_df = df.style \
        .format({'ai_performance': '{:.2f}'}, thousands=",", decimal=",")

    # Display the formatted DataFrame
    display(formatted_df)

```

```
except Exception as e:
    print(e)
```

	continent	income_groups	ai_performance
0	None	None	23,95
1	americas	high_income	44,20
2	asia	high_income	29,69
3	asia	upper_middle_income	30,37
4	europe	high_income	27,37

In [29]:

```
query = """
SELECT world_4region AS continent, income_groups, AVG(overall) AS ai_performance
FROM {view_name}
WHERE overall IS NOT NULL
GROUP BY continent, income_groups
HAVING ai_performance <= 20
ORDER BY continent, income_groups, ai_performance DESC
"""

try:
    # Execute the query and store the result in a DataFrame
    df = pd.read_sql_query(query, engine)

    # Apply formatting for thousands and decimal points
    formatted_df = df.style \
        .format({'ai_performance': '{:.2f}'}, thousands=",", decimal=",")
    # Display the formatted DataFrame
    display(formatted_df)
except Exception as e:
    print(e)
```

	continent	income_groups	ai_performance
0	africa	lower_middle_income	12,36
1	africa	upper_middle_income	14,10
2	americas	upper_middle_income	18,58
3	asia	lower_middle_income	19,42
4	europe	upper_middle_income	19,60

**Query 17 & 18 Results:** The conclusion from these two queries is that there is only one exception to the rule that you need to have a high income status to be a prominent ai region. That exception is of course China, which, although it is only in the upper middle income group, has such a large economy that it unavoidably plays a big role in the ai development scene. Investigation into subregions within China (Sjanghai, Beijing, ...) and the two semi-independent sub-regions of China (Hong Kong and Taiwan) might reveal that, just as in the division of the Americas, the difference in income groups is also quite relevant. This again would lead us too far.

## 5.4. Global Organizations and Networks

**Query 19:** Let's now have a look at organisational membership, a category that is murkier than income status. There are two different places where we can look at this data: our basic view and the organisation and membership tables from the mondial database. We want to combine both.

```
In [30]: query = f"""
SELECT club, COUNT(Code), AVG(overall) AS ai_performance
FROM {view_name}
WHERE overall IS NOT NULL
GROUP BY club
ORDER BY ai_performance DESC
"""

try:
    # Execute the query and store the result in a DataFrame
    df = pd.read_sql_query(query, engine)

    # Apply formatting for thousands and decimal points
    formatted_df = df.style \
        .format({'ai_performance': '{:.2f}'}, thousands=",", decimal=",")

    # Display the formatted DataFrame
    display(formatted_df)
except Exception as e:
    print(e)
```

	club	COUNT(Code)	ai_performance
0	oecd	31	31,17
1	None	2	23,95
2	others	6	21,30
3	g77	23	20,38

**Query 19 Results:** From this simple query, it is quite clear that OECD members are the best performers. For the other groups, there is no significant difference.

**Query 20:** Let's try to have a more detailed look at membership of international organisation to see if it tells us anything else.

*Note: For the report we have limited the output of the following query to 100 records. The original query contains 3223 rows which is not very nice reproduce in a pdf or printed report.*

```
In [31]: query = f"""
SELECT t1.club, t1.Code, t1.name, t1.overall, t2.organization, t2.type, t3.country AS HQ_Country, t4.name
FROM {view_name} AS t1
LEFT JOIN ismember AS t2 ON t1.code = t2.country
LEFT JOIN organization AS t3 ON t2.organization = t3.abbreviation
LEFT JOIN {view_name} AS t4 ON t3.country = t4.code
WHERE t1.overall IS NOT NULL
ORDER BY t1.overall DESC
LIMIT 100
"""

try:
    # Execute the query and store the result in a DataFrame
    df = pd.read_sql_query(query, engine)
```

```
# Apply formatting for thousands and decimal points
formatted_df = df.style \
    .format({'ai_performance': '{:.2f}'}, thousands=",", decimal=",")
# Display the formatted DataFrame
display(formatted_df)
except Exception as e:
    print(e)
```

	club	code	name	overall	organization	type	HQ_Country	name
0	oecd	USA	United States	100,000000	AfDB	nonregional member	CI	None
1	oecd	USA	United States	100,000000	AG	observer	PE	Peru
2	oecd	USA	United States	100,000000	ANC	member	B	Belgium
3	oecd	USA	United States	100,000000	ANZUS	member	AUS	Australia
4	oecd	USA	United States	100,000000	APEC	member	SGP	Singapore
5	oecd	USA	United States	100,000000	AsDB	nonregional member	RP	Philippines
6	oecd	USA	United States	100,000000	BIS	member	CH	Switzerland
7	oecd	USA	United States	100,000000	CCC	member	B	Belgium
8	oecd	USA	United States	100,000000	CP	member	CL	Sri Lanka
9	oecd	USA	United States	100,000000	EBRD	member	GB	United Kingdom
10	oecd	USA	United States	100,000000	ECE	member	CH	Switzerland
11	oecd	USA	United States	100,000000	ECLAC	member	RCH	Chile
12	oecd	USA	United States	100,000000	EN	member	None	None
13	oecd	USA	United States	100,000000	ESCAP	member	THA	Thailand
14	oecd	USA	United States	100,000000	FAO	member	I	Italy
15	oecd	USA	United States	100,000000	G-10	member	F	France
16	oecd	USA	United States	100,000000	G-2	member	None	None
17	oecd	USA	United States	100,000000	G-5	member	None	None
18	oecd	USA	United States	100,000000	G-7	member	None	None
19	oecd	USA	United States	100,000000	G-8	member	None	None
20	oecd	USA	United States	100,000000	IADB	member	USA	United States
21	oecd	USA	United States	100,000000	IAEA	member	A	Austria
22	oecd	USA	United States	100,000000	IBRD	member	USA	United States
23	oecd	USA	United States	100,000000	ICAO	member	CDN	Canada
24	oecd	USA	United States	100,000000	ICC	member	F	France
25	oecd	USA	United States	100,000000	ICFTU	member	B	Belgium
26	oecd	USA	United States	100,000000	ICRM	National Society	CH	Switzerland
27	oecd	USA	United States	100,000000	IDA	Part I	USA	United States
28	oecd	USA	United States	100,000000	IEA	member	F	France
29	oecd	USA	United States	100,000000	IFAD	Category I	I	Italy
30	oecd	USA	United States	100,000000	IFC	member	USA	United States

	club	code	name	overall	organization	type	HQ_Country	name
31	oecd	USA	United States	100,000000	IFRCS	member	CH	Switzerland
32	oecd	USA	United States	100,000000	ILO	member	CH	Switzerland
33	oecd	USA	United States	100,000000	IMF	member	USA	United States
34	oecd	USA	United States	100,000000	IMO	member	GB	United Kingdom
35	oecd	USA	United States	100,000000	Inmarsat	member	GB	United Kingdom
36	oecd	USA	United States	100,000000	Intelsat	member	USA	United States
37	oecd	USA	United States	100,000000	Interpol	member	F	France
38	oecd	USA	United States	100,000000	IOC National Olympic Committee	CH	Switzerland	
39	oecd	USA	United States	100,000000	IOM	member	CH	Switzerland
40	oecd	USA	United States	100,000000	ISO	member	CH	Switzerland
41	oecd	USA	United States	100,000000	ITU	member	CH	Switzerland
42	oecd	USA	United States	100,000000	MINURSO	member	USA	United States
43	oecd	USA	United States	100,000000	MTCR	member	None	None
44	oecd	USA	United States	100,000000	NATO	member	B	Belgium
45	oecd	USA	United States	100,000000	NSG	member	A	Austria
46	oecd	USA	United States	100,000000	OAS	member	USA	United States
47	oecd	USA	United States	100,000000	OECD	member	F	France
48	oecd	USA	United States	100,000000	OSCE	member	A	Austria
49	oecd	USA	United States	100,000000	PCA	member	None	None
50	oecd	USA	United States	100,000000	SPC	member	NCA	None
51	oecd	USA	United States	100,000000	UN	member	USA	United States
52	oecd	USA	United States	100,000000	UNCRO	member	USA	United States
53	oecd	USA	United States	100,000000	UNHCR	member	CH	Switzerland
54	oecd	USA	United States	100,000000	UNIDO	member	A	Austria
55	oecd	USA	United States	100,000000	UNIKOM	member	USA	United States
56	oecd	USA	United States	100,000000	UNITAR	member	None	None
57	oecd	USA	United States	100,000000	UNMIH	member	USA	United States
58	oecd	USA	United States	100,000000	UNOMIG	member	USA	United States
59	oecd	USA	United States	100,000000	UNPREDEP	member	None	None
60	oecd	USA	United States	100,000000	UNPROFOR	member	USA	United States
61	oecd	USA	United States	100,000000	UNRWA	member	A	Austria

	club	code	name	overall	organization	type	HQ_Country	name
62	oecd	USA	United States	100,000000	UNTSO	member	IL	Israel
63	oecd	USA	United States	100,000000	UNU	member	None	None
64	oecd	USA	United States	100,000000	UPU	member	CH	Switzerland
65	oecd	USA	United States	100,000000	WCL	member	B	Belgium
66	oecd	USA	United States	100,000000	WHO	member	CH	Switzerland
67	oecd	USA	United States	100,000000	WIPO	member	CH	Switzerland
68	oecd	USA	United States	100,000000	WMO	member	CH	Switzerland
69	oecd	USA	United States	100,000000	WToO	member	E	Spain
70	oecd	USA	United States	100,000000	WTrO	member	None	None
71	oecd	USA	United States	100,000000	ZC	member	None	None
72	g77	TJ	China	61,500000	AfDB	nonregional member	CI	None
73	g77	TJ	China	61,500000	APEC	member	SGP	Singapore
74	g77	TJ	China	61,500000	AsDB	regional member	RP	Philippines
75	g77	TJ	China	61,500000	CCC	member	B	Belgium
76	g77	TJ	China	61,500000	ESCAP	member	THA	Thailand
77	g77	TJ	China	61,500000	FAO	member	I	Italy
78	g77	TJ	China	61,500000	IAEA	member	A	Austria
79	g77	TJ	China	61,500000	IBRD	member	USA	United States
80	g77	TJ	China	61,500000	ICAO	member	CDN	Canada
81	g77	TJ	China	61,500000	ICFTU	member	B	Belgium
82	g77	TJ	China	61,500000	ICRM	National Society	CH	Switzerland
83	g77	TJ	China	61,500000	IDA	Part II	USA	United States
84	g77	TJ	China	61,500000	IFAD	Category III	I	Italy
85	g77	TJ	China	61,500000	IFC	member	USA	United States
86	g77	TJ	China	61,500000	IFRCS	member	CH	Switzerland
87	g77	TJ	China	61,500000	ILO	member	CH	Switzerland
88	g77	TJ	China	61,500000	IMF	member	USA	United States
89	g77	TJ	China	61,500000	IMO	member	GB	United Kingdom
90	g77	TJ	China	61,500000	Inmarsat	member	GB	United Kingdom
91	g77	TJ	China	61,500000	Intelsat	member	USA	United States
92	g77	TJ	China	61,500000	Interpol	member	F	France

	club	code	name	overall	organization		type	HQ_Country		name
93	g77	TJ	China	61,500000	IOC	National Olympic Committee		CH	Switzerland	
94	g77	TJ	China	61,500000	ISO		member	CH	Switzerland	
95	g77	TJ	China	61,500000	ITU		member	CH	Switzerland	
96	g77	TJ	China	61,500000	LAIA		observer	ROU	Uruguay	
97	g77	TJ	China	61,500000	Mekong Group		member	None	None	
98	g77	TJ	China	61,500000	MINURSO		member	USA	United States	
99	g77	TJ	China	61,500000	NAM		observer	RI	Indonesia	

**Query 20 Results:** That's way to much information to draw decent conclusions.

**Query 21:** Therefore, let's count the number of memberships the countries have and see if we can draw conclusion on that basis.

```
In [32]: query = """
SELECT t1.club, t1.Code, t1.name, t1.overall, count(t2.organization) AS Mbr_Count
FROM {view_name} AS t1
LEFT JOIN ismember AS t2 ON t1.code = t2.country
LEFT JOIN organization AS t3 ON t2.organization = t3.abbreviation
LEFT JOIN {view_name} AS t4 ON t3.country = t4.code
WHERE t1.overall IS NOT NULL
GROUP BY t1.club, t1.Code, t1.name, t1.overall
ORDER BY Mbr_Count DESC
"""

try:
    # Execute the query and store the result in a DataFrame
    df = pd.read_sql_query(query, engine)

    # Apply formatting for thousands and decimal points
    formatted_df = df.style \
        .format({
            'ai_performance': '{:.2f}',
            'overall': '{:.2f}',
            'thousands': '.', 'decimal': ','})
    # Display the formatted DataFrame
    display(formatted_df)
except Exception as e:
    print(e)
```

	club	code	name	overall	Mbr_Count
0	oecd	F	France	32,80	83
1	oecd	GB	United Kingdom	41,80	74
2	oecd	CDN	Canada	40,30	73
3	oecd	USA	United States	100,00	72
4	oecd	S	Sweden	30,30	71
5	oecd	I	Italy	26,50	71
6	oecd	SF	Finland	34,90	70
7	oecd	D	Germany	39,20	70
8	oecd	NL	Netherlands	34,50	70
9	oecd	A	Austria	27,70	69
10	oecd	B	Belgium	26,60	68
11	oecd	DK	Denmark	30,50	67
12	oecd	E	Spain	27,70	65
13	g77	RA	Argentina	17,50	64
14	oecd	CH	Switzerland	37,70	64
15	oecd	N	Norway	26,40	64
16	g77	ET	Egypt	16,90	63
17	oecd	P	Portugal	23,70	61
18	oecd	PL	Poland	24,80	59
19	g77	PK	Pakistan	10,10	59
20	oecd	MEX	Mexico	16,90	58
21	g77	BR	Brazil	22,10	58
22	oecd	GR	Greece	18,30	58
23	oecd	J	Japan	33,90	58
24	oecd	TR	Turkey	20,60	56
25	others	R	Russia	23,70	56
26	oecd	AUS	Australia	30,90	55
27	g77	IND	India	31,40	55
28	oecd	H	Hungary	20,70	54
29	g77	WAN	Nigeria	9,30	52
30	g77	ROU	Uruguay	16,30	51

	club	code	name	overall	Mbr_Count
<b>31</b>	oecd	IRL	Ireland	28,80	51
<b>32</b>	g77	TN	Tunisia	13,70	51
<b>33</b>	oecd	NZ	New Zealand	21,60	50
<b>34</b>	g77	RI	Indonesia	18,20	50
<b>35</b>	g77	CO	Colombia	17,80	50
<b>36</b>	g77	MAL	Malaysia	19,60	49
<b>37</b>	oecd	CZ	Czech Republic	22,10	49
<b>38</b>	g77	RCH	Chile	20,20	48
<b>39</b>	g77	EAK	Kenya	8,30	47
<b>40</b>	oecd	L	Luxembourg	29,20	47
<b>41</b>	g77	MA	Morocco	13,60	46
<b>42</b>	oecd	SK	Slovakia	17,10	46
<b>43</b>	g77	SA	Saudi Arabia	23,30	44
<b>44</b>	oecd	ROK	South Korea	40,30	44
<b>45</b>	g77	TJ	China	61,50	44
<b>46</b>	oecd	IS	Iceland	20,60	43
<b>47</b>	g77	CL	Sri Lanka	10,00	43
<b>48</b>	g77	UAE	United Arab Emirates	23,90	41
<b>49</b>	oecd	IL	Israel	40,00	39
<b>50</b>	others	M	Malta	22,40	39
<b>51</b>	g77	RSA	South Africa	14,10	38
<b>52</b>	g77	Q	Qatar	19,80	37
<b>53</b>	others	SLO	Slovenia	21,50	37
<b>54</b>	g77	VN	Vietnam	18,00	37
<b>55</b>	g77	SGP	Singapore	49,70	34
<b>56</b>	others	EW	Estonia	26,00	34
<b>57</b>	g77	BRN	Bahrain	13,50	34
<b>58</b>	others	ARM	Armenia	14,50	33
<b>59</b>	others	LT	Lithuania	19,70	33
<b>60</b>	None	HONX	Hong Kong	22,50	11
<b>61</b>	None	RC	Taiwan	25,40	7

**Query 21 Results:** Still, not a very clear picture.

**Query 22:** We could go back to our view of economic rankings and see whether there is a correlation between these rankings and the number of HQs that each organization has in a country. We only select countries with a decent GDP, GDP per capita and a score in the ranking table.

For analysis purposes, we also calculated the correlation coefficient and visualized the results using a linear regression plot.

```
In [33]: query = f"""
SELECT t1.club, t1.Code, t1.name, t1.overall, t2.overall_rank, count(t4.abbreviation) AS HQ_Count
FROM {view_name} AS t1
LEFT JOIN {view_name_economic} AS t2 ON t1.name = t2.name
LEFT JOIN ismember AS t3 ON t1.code = t3.country
LEFT JOIN organization AS t4 ON t3.organization = t4.abbreviation
LEFT JOIN {view_name} AS t5 ON t4.country = t5.code
WHERE t2.overall_rank < 63 AND t2.GDP_rank <= 100 AND t2.GDPPC_rank <= 100 AND t1.code = t5.code
GROUP BY t1.club, t1.Code, t1.name, t1.overall
ORDER BY t2.overall_rank
"""

try:
    # Apply formatting using the modified function
    df = pd.read_sql_query(query, engine)
    formatted_df = df.style\
        .format(precision=0, thousands=",", decimal=",")
    # Display the formatted DataFrame
    display(formatted_df)
except Exception as e:
    print(e)
```

	club	code	name	overall	overall_rank	HQ_Count
0	oecd	USA	United States	100	1	14
1	oecd	GB	United Kingdom	42	4	4
2	oecd	CDN	Canada	40	5	1
3	oecd	CH	Switzerland	38	9	17
4	oecd	SF	Finland	35	10	1
5	oecd	F	France	33	13	10
6	oecd	AUS	Australia	31	15	1
7	oecd	S	Sweden	30	17	1
8	oecd	E	Spain	28	20	1
9	oecd	A	Austria	28	20	4
10	oecd	B	Belgium	27	22	8
11	oecd	I	Italy	26	23	2
12	g77	UAE	United Arab Emirates	24	28	1
13	g77	SA	Saudi Arabia	23	31	3
14	oecd	CZ	Czech Republic	22	34	1
15	oecd	H	Hungary	21	38	1
16	g77	RCH	Chile	20	41	1
17	g77	RI	Indonesia	18	46	2
18	g77	RA	Argentina	18	49	1
19	oecd	MEX	Mexico	17	51	1
20	g77	ROU	Uruguay	16	53	1
21	g77	RSA	South Africa	14	55	1
22	g77	MA	Morocco	14	57	1
23	g77	CL	Sri Lanka	10	60	1

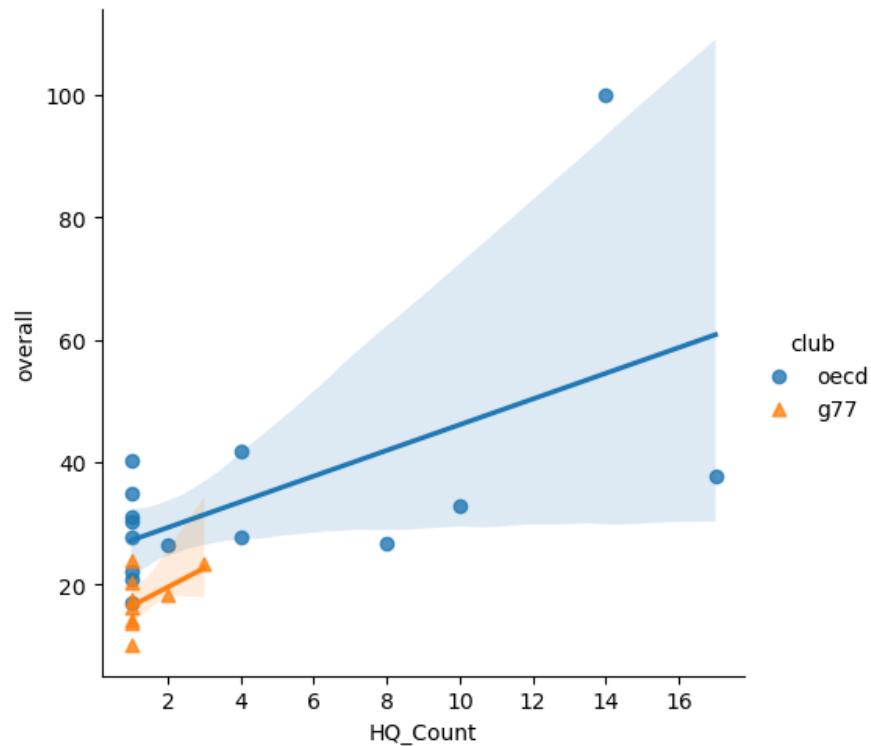
```
In [34]: corrCo = df['overall'].corr(df['HQ_Count'])
print("Correlation coefficient: " + str(f'{corrCo:.2f}'))
```

Correlation coefficient: 0.64

```
In [35]: sns.lmplot(y='overall', x='HQ_Count', hue='club', markers=['o', '^'], data=df)
```

```
C:\Users\geert\anaconda3\lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

```
Out[35]: <seaborn.axisgrid.FacetGrid at 0x1c5ab152150>
```



**Query 22 Results:** Again, there are top ranking countries on the ai arena that are not prominent on the international stage. A correlation coefficient of 64% isn't really impressive, however also not neglectable. Most interesting is that we see a similar slope and distribution pattern between different "clubs" in the linear regression plot.

An explanation for the limited correlation could be due to the available data or to the fact that the international organisations mentioned in this list are not so relevant for this domain. You could say that that also means that some prominent countries on the international stage have little influence on the ai developments currently going on.

Here as well as in the next queries, the most prominent absent country is China. This probably means that the data from the organization database are not quite complete. It is however a fact that the prominence of China on the international stage is relatively recent and the establishment of most organization date from a time before that period.

**Query 23 & 24:** As a last check we can take another table that tells something about the connectedness of countries specifically concerning research.

First thing we can do is check how many countries have been included in this table. The next query takes us a bit further and makes the link to our basic\_country\_data view and sums the value of the connections between countries that are in our view, giving a net value of the found relationships for each country, which are visualized in horizontal bar chart.

To make things a bit clearer we again calculated the correlation coefficient between the netvalue and the overall ai score and visualized this score in a regression plot.

```
In [36]: # Counting distinct countries ...
query =f'''SELECT COUNT(DISTINCT t1.`First Country`)
FROM (
# ... that are also in our basic_country_data
    SELECT t2.`First Country`
    FROM ai_oecd_research_networks AS t2
```

```

    INNER JOIN {view_name} AS t3 ON t2.`First Country` = t3.name
) AS t1
"""

try:
    # Apply formatting using the modified function
    df = pd.read_sql_query(query, engine)
    formatted_df = df.style\
        .format(precision=0, thousands=",", decimal=",")
    # Display the formatted DataFrame
    display(formatted_df)
except Exception as e:
    print(e)

```

**COUNT(DISTINCT t1.`First Country`)**

0	163
---	-----

In [37]: # Selecting all countries in the network table with ai scores and summed network values

```

query ="""
SELECT t1.`First Country` AS Country_Net, t2.overall, round(sum(t1.value),2) AS NetValue, t2.club
FROM ai_oecd_research_networks AS t1
LEFT JOIN {view_name} AS t2 ON t1.`First Country` = t2.name
# ... that are in our basic_country_data
WHERE t2.overall IS NOT NULL
GROUP BY t1.`First Country`, t2.overall
# ... using only those countries that have a significant network value
HAVING NetValue > 1000
ORDER BY t2.overall DESC
"""

try:
    # Apply formatting using the modified function
    df = pd.read_sql_query(query, engine)
    formatted_df = df.style\
        .format(precision=0, thousands=",", decimal=",")
    # Display the formatted DataFrame
    display(formatted_df)
except Exception as e:
    print(e)

```

	Country_Net	overall	NetValue	club
0	United States	100	264.389	oecd
1	China	62	87.504	g77
2	Singapore	50	10.598	g77
3	United Kingdom	42	81.012	oecd
4	Canada	40	44.247	oecd
5	Israel	40	9.784	oecd
6	Germany	39	61.130	oecd
7	Switzerland	38	16.932	oecd
8	Finland	35	7.148	oecd
9	Netherlands	34	26.140	oecd
10	Japan	34	59.409	oecd
11	France	33	74.411	oecd
12	India	31	12.574	g77
13	Australia	31	31.175	oecd
14	Denmark	30	7.585	oecd
15	Sweden	30	13.743	oecd
16	Ireland	29	4.906	oecd
17	Austria	28	8.801	oecd
18	Spain	28	29.759	oecd
19	Belgium	27	13.854	oecd
20	Italy	26	40.228	oecd
21	Norway	26	6.595	oecd
22	Taiwan	25	18.066	None
23	Poland	25	11.317	oecd
24	United Arab Emirates	24	1.029	g77
25	Portugal	24	7.625	oecd
26	Russia	24	11.409	others
27	Saudi Arabia	23	1.044	g77
28	Brazil	22	13.883	g77
29	Czech Republic	22	5.737	oecd
30	New Zealand	22	4.803	oecd

	Country_Net	overall	NetValue	club
31	Slovenia	22	2.217	others
32	Hungary	21	5.153	oecd
33	Turkey	21	6.686	oecd
34	Chile	20	2.718	g77
35	Malaysia	20	3.071	g77
36	Greece	18	8.689	oecd
37	Colombia	18	1.477	g77
38	Argentina	18	3.719	g77
39	Mexico	17	7.451	oecd
40	Egypt	17	2.903	g77
41	South Africa	14	2.634	g77
42	Tunisia	14	1.190	g77
43	Morocco	14	1.138	g77
44	Pakistan	10	1.261	g77

In [38]:

```
import matplotlib.pyplot as plt

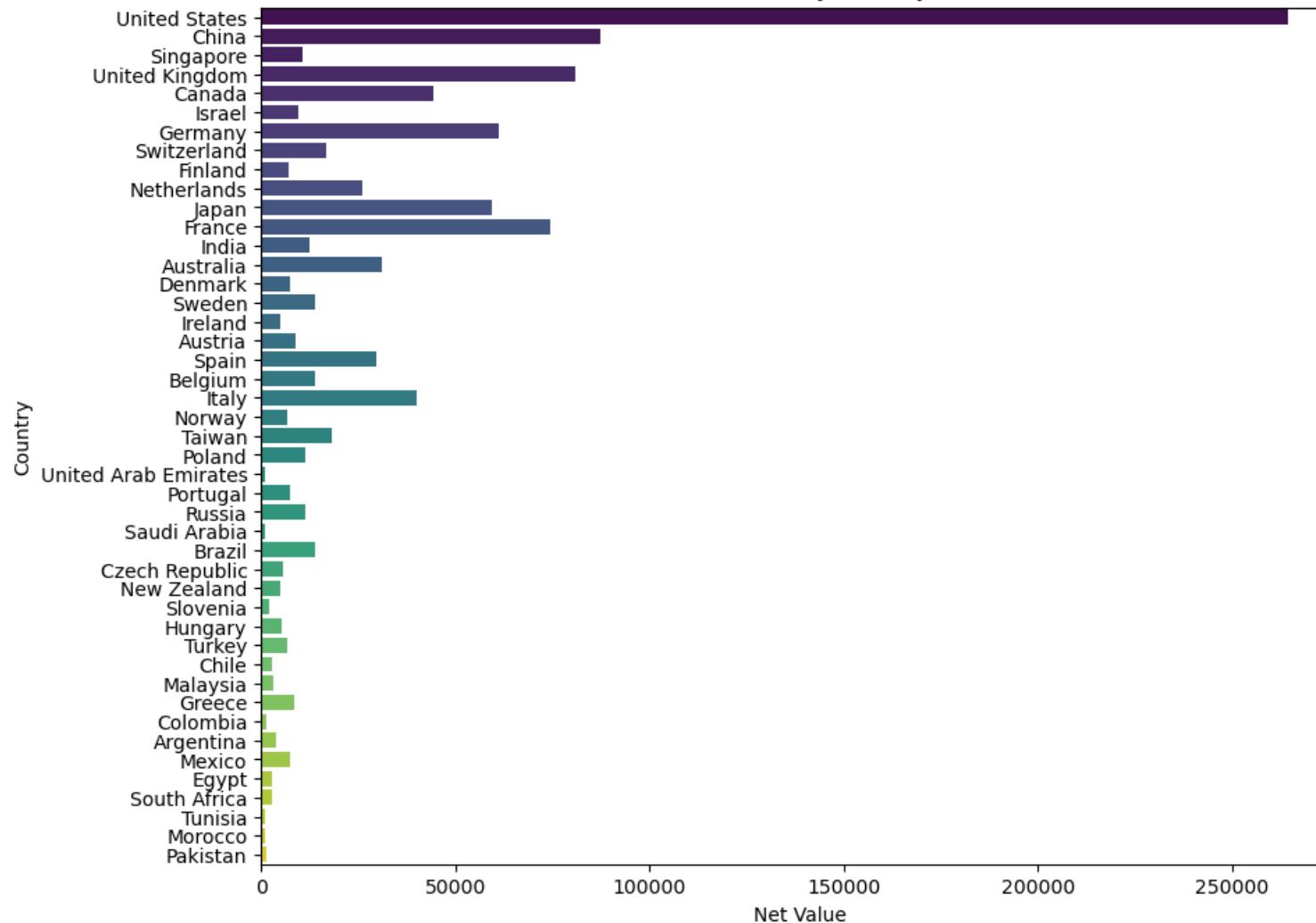
# Ensure that the data is sorted in descending order for better visualization
df_sorted = df.sort_values('overall', ascending=False)

# Create the plot
plt.figure(figsize=(10, 8)) # Adjust the size as needed
sns.barplot(x='NetValue', y='Country_Net', data=df_sorted, palette='viridis')

# Adding labels and title for clarity
plt.xlabel('Net Value')
plt.ylabel('Country')
plt.title('Net Values by Country')

# Show the plot
plt.show()
```

Net Values by Country



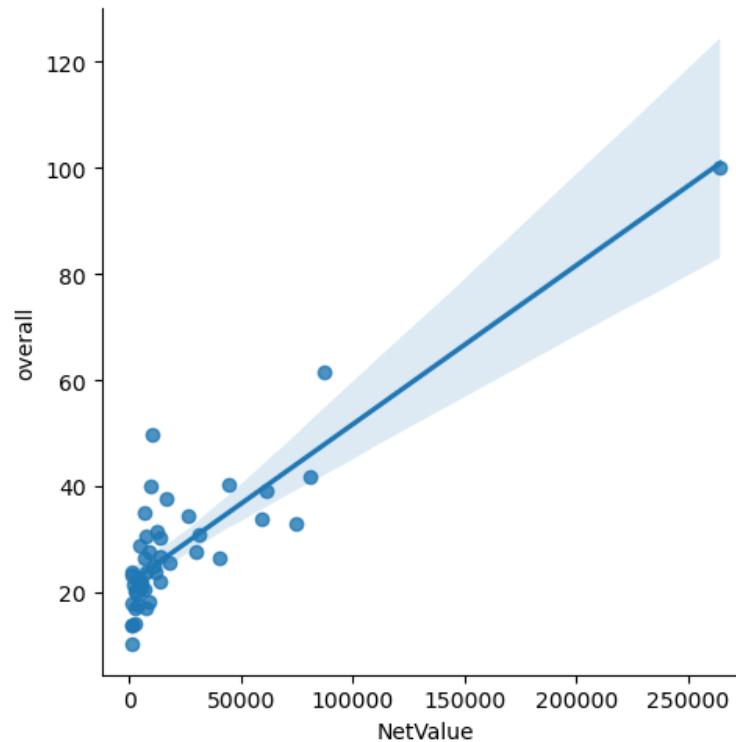
```
In [39]: corrCo = df['overall'].corr(df['NetValue'])
print("Correlation coefficient: " + str(f'{corrCo:.2f}'))
```

Correlation coefficient: 0.87

```
In [40]: sns.lmplot(y='overall', x='NetValue', data=df)
```

```
C:\Users\geert\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The figure layout has changed to tight
    self._figure.tight_layout(*args, **kwargs)
<seaborn.axisgrid.FacetGrid at 0x1c5ab46eb90>
```

Out[40]:



**Query 23 & 24 Results:** This list is somewhat clearer. There seems to be a large correlation (87%) between research network values and prominence on the ai stage. However there are some notable exceptions. These exceptions are again interesting to investigate. We can also see that some countries that are in the basic\_country\_table are missing in the network relations table.

The list of countries with prominent research networks over time but significant lower scores on the ai ranking are France, Japan, Italy and Spain. On the other hand, certain countries score prominently in the ai index and have low research network scores, such as Singapore, Israel and Finland. This could mean that the latter countries primarily perform their AI research within their own country while the former countries might have networks not fully deployed or underperforming in their impact. Again, further research is necessarily to answer such questions.

## 6. Conclusions

There are many more possibilities to investigate why certain countries have prominence in the ongoing ai revolution, however, within the confines of this project, we will limit ourselves to some of the hints for further investigations we referred to earlier. Many of the tables we collected to support such an investigation were not used but could be interesting to look at. It is also very important to update the data from the mondial database to make accurate conclusions as to which factors do count for ai prominence as a country. As said, all this would take much more research, but it would definitely be worthwhile.