Data and SQL
GLOBAL HEALTH DATABASE

AVRIL CHILDS

- Create new database Global_Health which will contain a database of health data.

- For purposes of this project, this will only contain a list of 20 countries

- Data retrieved from https://ourworldindata.org

TABLE 1: World Health Organisation (WHO) Regions

WHO_Regions: This table contains information about different regions defined by the World Health Organization (WHO). It includes the region ID, region name, and total population in millions for each region.

```sql
-- Database of health data of different countries around the world
CREATE DATABASE Global_Health;

USE Global_Health;

CREATE TABLE WHO_Regions (
    Region_ID CHAR(2) NOT NULL Primary Key,
    REGION ENUM
        ("Africa", "Americas", "South-East Asia",
        "Europe", "Eastern Mediterranean", "Western Pacific"),
    Total_Population_mil INT
);

INSERT INTO WHO_REGIONS (Region_ID, REGION, Total_Population_mil)
VALUE
("R1", "Africa", 1163),
("R2", "Americas", 1030),
("R3", "South-East Asia", 2056),
("R4", "Europe", 931),
("R5", "Eastern Mediterranean", 767),
("R6", "Western Pacific", 1933);

SELECT * from WHO_REGIONS;
```

Region_ID is Primary Key. ENUM used to specify entry options

	Region_ID	REGION	Total_Population_mil
▶	R1	Africa	1163
	R2	Americas	1030
	R3	South-East Asia	2056
	R4	Europe	931
	R5	Eastern Mediterranean	767
	R6	Western Pacific	1933
*	NULL	NULL	NULL

TABLE 1 – WHO_REGIONS

```sql
-- Table 2 - Countries representated in this database

CREATE TABLE Representative_Countries (
Country_ID VARCHAR(3) NOT NULL Primary key,
Country_name VARCHAR (50),
Region_ID CHAR(2) NOT NULL,
Foreign Key (Region_ID) References WHO_REGIONS(REGION_ID)
);

INSERT INTO Representative_Countries
(Country_ID, country_name, region_ID)
Value
("C01", "Afghanistan", "R5"),
("C02", "Argentina", "R2"),
("C03", "Australia", "R6"),
("C04", "Bangladesh", "R3"),
("C05", "Brazil", "R2"),
("C06", "Canada", "R2"),
("C07", "China", "R6"),
("C08", "Ethiopia", "R1"),
("C09", "Finland", "R4"),
("C10", "France", "R4"),
("C11", "Germany", "R4"),
("C12", "Ghana", "R1"),
("C13", "Indonesia", "R3"),
("C14", "India", "R3"),
("C15", "Japan", "R6"),
("C16", "Russia", "R4"),
("C17", "Singapore", "R6"),
("C18", "Somalia", "R5"),
("C19", "UK", "R4"),
("C20", "USA", "R2");
```

Foreign key to link with Table 1 Region_ID

Table 2: Representative Countries, with WHO Region

Representative_Countries: This table represents the countries included in the database and their respective regions. It has a foreign key reference to the WHO_Regions table.

Table 2 Representative_Countries

Country_ID	Country_name	Region_ID
C01	Afghanistan	R5
C02	Argentina	R2
C03	Australia	R6
C04	Bangladesh	R3
C05	Brazil	R2
C06	Canada	R2
C07	China	R6
C08	Ethiopia	R1
C09	Finland	R4
C10	France	R4
C11	Germany	R4
C12	Ghana	R1
C13	Indonesia	R3
C14	India	R3
C15	Japan	R6
C16	Russia	R4
C17	Singapore	R6
C18	Somalia	R5
C19	UK	R4
C20	USA	R2
NULL	NULL	NULL

```sql
-- Create Table 3 - County Health expenditures and Population
CREATE Table Country_Healthex_Pop (
    Country_ID VARCHAR(3) NOT NULL Primary key,
    Country_name VARCHAR (50) NOT NULL,
    Health_Expenditure DEC(10,2), -- 10 digits, 2 decimal points,
    Population_per_million DEC(6,2),
    FOREIGN KEY (Country_Id) REFERENCES Representative_Countries(Country_Id)
);
```

```sql
INSERT INTO Country_Healthex_Pop
(Country_ID, country_name, Health_Expenditure, Population_per_million)
Value
("C01", "Afghanistan", 285.56, 37.77),
("C02", "Argentina", 2198.88, 44.75),
("C03", "Australia", 5294.46, 25.36),
("C04", "Bangladesh", 123.29, 165.52),
("C05", "Brazil", 1497.81, 211.78),
("C06", "Canada", 5520.65, 37.52),
("C07", "China", 880.19, 1421.86),
("C08", "Ethiopia", 75.11, 114.12),
("C09", "Finland", 4710.00, 5.52),
("C10", "France", 5492.53, 64.40),
("C11", "Germany", 6738.67, 83.15),
("C12", "Ghana", 193.22, 31.52),
("C13", "Indonesia", 358.29, 269.58),
("C14", "India", 211, 1383.11),
("C15", "Japan", 4587.03, 125.79),
("C16", "Russia", 1704.04, 145.74),
("C17", "Singapore", 4102.27, 5.87),
("C18", "Somalia", NULL, 15.99),
("C19", "UK", 5087.38, 66.78),
("C20", "USA", 10921.01, 334.32);
```

Foreign Key linked to table 2 for Country ID

Table 3: Health Expenditure ($ per capita) and Population (million) by Country

TABLE 3 - COUNTRY_HEALTHEX_POP

Country_ID	Country_name	Health_Expenditure	Population_per_r
C01	Afghanistan	285.56	37.77
C02	Argentina	2198.88	44.75
C03	Australia	5294.46	25.36
C04	Bangladesh	123.29	165.52
C05	Brazil	1497.81	211.78
C06	Canada	5520.65	37.52
C07	China	880.19	1421.86
C08	Ethiopia	75.11	114.12
C09	Finland	4710.00	5.52
C10	France	5492.53	64.40
C11	Germany	6738.67	83.15
C12	Ghana	193.22	31.52
C13	Indonesia	358.29	269.58
C14	India	211.00	1383.11
C15	Japan	4587.03	125.79
C16	Russia	1704.04	145.74
C17	Singapore	4102.27	5.87
C18	Somalia	NULL	15.99
C19	UK	5087.38	66.78
C20	USA	10921.01	334.32
NULL	NULL	NULL	NULL

```sql
-- Table 4 - Life expectancy of males and females from 2019 data
CREATE TABLE Life_Expectancy (
Country_Id VARCHAR (3) NOT NULL,
Country_name varchar(50) NOT NULL,
Male_Life_Ex DEC(5,2),
Female_Life_Ex DEC(5,2),
FOREIGN KEY (Country_Id) REFERENCES Representative_Countries(Country_Id)
);

INSERT INTO Life_expectancy
(Country_id, Country_name, Male_Life_Ex, Female_Life_Ex)
Values
("C01", "Afghanistan", 60.60, 66.70),
("C02", "Argentina", 73.90, 80.70),
("C03", "Australia", 81.10, 85.10),
("C04", "Bangladesh", 70.70, 75.10),
("C05", "Brazil", 72.20, 78.50),
("C06", "Canada", 80.30, 84.40),
("C07", "China", 75.30, 80.80),
("C08", "Ethiopia", 63.10, 68.80),
("C09", "Finland", 79.20, 84.50),
("C10", "France", 79.70, 85.60),
("C11", "Germany", 79.10, 84.00),
("C12", "Ghana", 62.70, 66.80),
("C13", "Indonesia", 68.50, 72.60),
("C14", "India", 69.50, 72.40),
("C15", "Japan", 81.40, 87.40),
("C16", "Russia", 68.80, 78.80),
("C17", "Singapore", 81.60, 85.90),
("C18", "Somalia", 55.10, 59.10),
("C19", "UK", 79.90, 83.50),
("C20", "USA", 76.60, 81.70);
```

Table 4: Life Expectancy (in yrs)

```sql
-- Table 5 Infant and child mortality data 50 yr comparison
CREATE TABLE Child_Mortality (
Country_id VARCHAR(3) NOT NULL primary key,
Country_name VARCHAR(50) NOT NULL,
Mortality1969_pc DEC(4,2),
Mortality2019_pc DEC(4,2)
);

INSERT into Child_Mortality
(Country_id, Country_name, Mortality1969_pc, Mortality2019_pc)
Value
("C01", "Afghanistan", 30.79, 6.01),
("C02", "Argentina", 7.25, 0.91),
("C03", "Australia", 2.17, 0.37),
("C04", "Bangladesh", 22.87, 3.07),
("C05", "Brazil", 13.65, 1.49),
("C06", "Canada", 2.30, 0.51),
("C07", "China", 11.84, 0.79),
("C08", "Ethiopia", 24.76, 5.08),
("CO9", "Finland", 1.68, 0.23),
("C10", "France", 1.91, 0.43),
("C11", "Germany", 2.64, 0.37),
("C12", "Ghana", 20.30, 4.64),
("C13", "Indonesia", 17.19, 2.38),
("C14", "India", 21.65, 3.44),
("C15", "Japan", 1.85, 0.25),
("C16", "Russia", 3.78, 0.58),
("C17", "Singapore", 2.84, 0.24),
("C18", "Somalia", 19.88, 11.83),
("C19", "UK", 2.14, 0.43),
("C20", "USA", 2.41, 0.64);

SELECT * FROM Child_Mortality
```

Table 5: Child mortality (%), comparing 1969 and 2019

```sql
ALTER TABLE Child_mortality
ADD CONSTRAINT
fk_country_id
FOREIGN KEY
(country_id)
REFERENCES
representative_countries
(country_id);
```

Table 4: Male and Female Life Expectancy (in yrs)

Country_Id	Country_name	Male_Life_Ex	Female_Life_Ex
C01	Afghanistan	60.60	66.70
C02	Argentina	73.90	80.70
C03	Australia	81.10	85.10
C04	Bangladesh	70.70	75.10
C05	Brazil	72.20	78.50
C06	Canada	80.30	84.40
C07	China	75.30	80.80
C08	Ethiopia	63.10	68.80
C09	Finland	79.20	84.50
C10	France	79.70	85.60
C11	Germany	79.10	84.00
C12	Ghana	62.70	66.80
C13	Indonesia	68.50	72.60
C14	India	69.50	72.40
C15	Japan	81.40	87.40
C16	Russia	68.80	78.80
C17	Singapore	81.60	85.90
C18	Somalia	55.10	59.10
C19	UK	79.90	83.50
C20	USA	76.60	81.70

Table 5: infant and Child Mortality (in % births)

Country_id	Country_name	Mortality1969_pc	Mortality2019_pc
C01	Afghanistan	30.79	6.01
C02	Argentina	7.25	0.91
C03	Australia	2.17	0.37
C04	Bangladesh	22.87	3.07
C05	Brazil	13.65	1.49
C06	Canada	2.30	0.51
C07	China	11.84	0.79
C08	Ethiopia	24.76	5.08
C10	France	1.91	0.43
C11	Germany	2.64	0.37
C12	Ghana	20.30	4.64
C13	Indonesia	17.19	2.38
C14	India	21.65	3.44
C15	Japan	1.85	0.25
C16	Russia	3.78	0.58
C17	Singapore	2.84	0.24
C18	Somalia	19.88	11.83
C19	UK	2.14	0.43
C20	USA	2.41	0.64
CO9	Finland	1.68	0.23
NULL	NULL	NULL	NULL

```sql
SELECT
    c.Country_ID, c.Country_name, r.Region,
    p.Health_Expenditure, p.Population_per_million
from Representative_countries c
LEFT JOIN WHO_Regions r ON c.Region_ID = r.region_ID
LEFT JOIN Country_Healthex_Pop p ON c.country_ID = p.country_ID
Order by p.Health_expenditure ASC;
```

Country_ID	Country_name	Region	Health_Expenditure	Population_per_million
C18	Somalia	Eastern Mediterranean	NULL	15.99
C08	Ethiopia	Africa	75.11	114.12
C04	Bangladesh	South-East Asia	123.29	165.52
C12	Ghana	Africa	193.22	31.52
C14	India	South-East Asia	211.00	1383.11
C01	Afghanistan	Eastern Mediterranean	285.56	37.77
C13	Indonesia	South-East Asia	358.29	269.58
C07	China	Western Pacific	880.19	1421.86
C05	Brazil	Americas	1497.81	211.78
C16	Russia	Europe	1704.04	145.74
C02	Argentina	Americas	2198.88	44.75
C17	Singapore	Western Pacific	4102.27	5.87
C15	Japan	Western Pacific	4587.03	125.79
C09	Finland	Europe	4710.00	5.52
C19	UK	Europe	5087.38	66.78
C03	Australia	Western Pacific	5294.46	25.36
C10	France	Europe	5492.53	64.40
C06	Canada	Americas	5520.65	37.52
C11	Germany	Europe	6738.67	83.15
C20	USA	Americas	10921.01	334.32

LEFT JOIN – COMBINING DATA FROM 3 TABLES

1. REPRESENTATIVE COUNTRIES,
2. WHO REGIONS
3. COUNTRY_HEALTHEX_POP

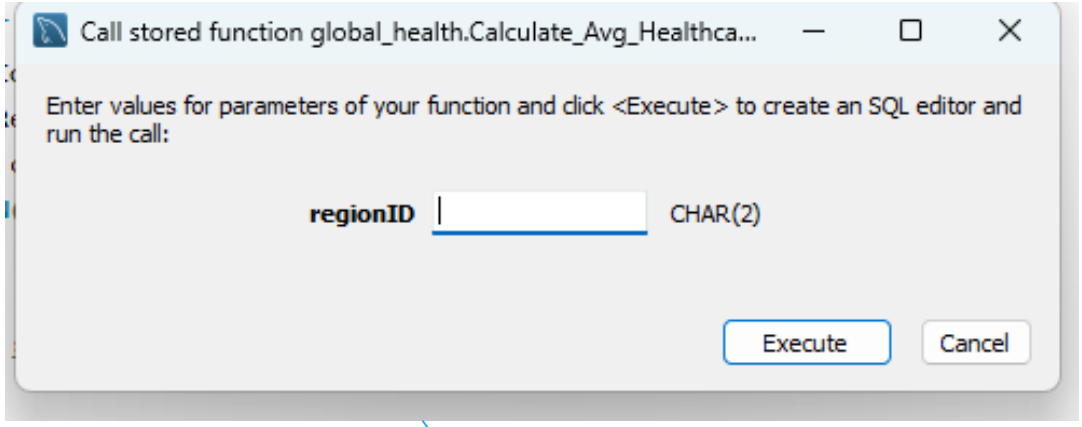
Stored Function

- Aim is to Calculate Average Healthcare Expenditure of each region when the function is called up.
- AVG function to calculate the average healthcare expenditure.
- JOIN links "Country_Health_Pop' table and 'Representative_Countries" table and "Where" filters the result by the "regionID" that the user keyed in.

```
190  DELIMITER //
191
192  CREATE FUNCTION Calculate_Avg_Healthcare_Expenditure (regionID CHAR(2))
193  RETURNS DECIMAL(10, 2)
194  DETERMINISTIC
195  BEGIN
196
197      DECLARE Avg_expenditure DECIMAL (10,2);
198      SELECT avg(Health_Expenditure) INTO avg_expenditure
199      FROM Country_Healthex_Pop p
200      JOIN Representative_Countries c ON p.Country_id = c.Country_id
201      WHERE c.Region_ID = regionID;
202      RETURN(avg_expenditure);
203  END
204  //
205  DELIMITER ;
206
```

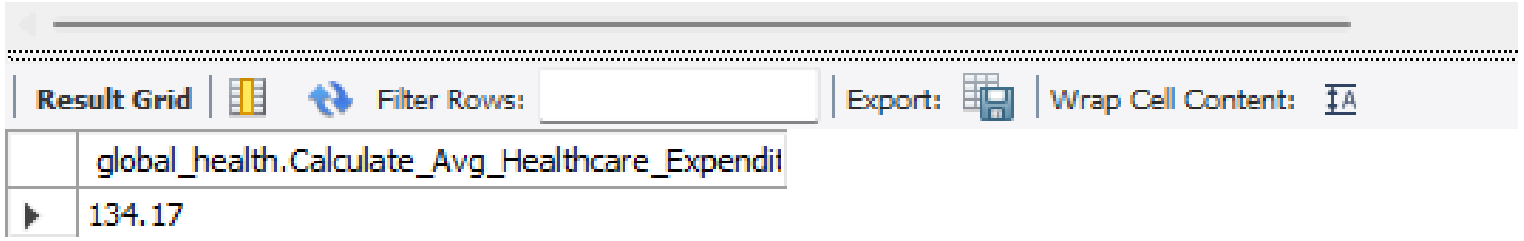
Country_ID	Country_name	Region_ID
C01	Afghanistan	R5
C02	Argentina	R2
C03	Australia	R6
C04	Bangladesh	R3
C05	Brazil	R2
C06	Canada	R2
C07	China	R6
C08	Ethiopia	R1
C09	Finland	R4
C10	France	R4
C11	Germany	R4
C12	Ghana	R1
C13	Indonesia	R3
C14	India	R3
C15	Japan	R6
C16	Russia	R4
C17	Singapore	R6
C18	Somalia	R5
C19	UK	R4
C20	USA	R2
NULL	NULL	NULL

Country_ID	Country_name	Health_Expenditure	Population_per_million
C01	Afghanistan	285.56	37.77
C02	Argentina	2198.88	44.75
C03	Australia	5294.46	25.36
C04	Bangladesh	123.29	165.52
C05	Brazil	1497.81	211.78
C06	Canada	5520.65	37.52
C07	China	880.19	1421.86
C08	Ethiopia	75.11	114.12
C09	Finland	4710.00	5.52
C10	France	5492.53	64.40
C11	Germany	6738.67	83.15
C12	Ghana	193.22	31.52
C13	Indonesia	358.29	269.58
C14	India	211.00	1383.11
C15	Japan	4587.03	125.79
C16	Russia	1704.04	145.74
C17	Singapore	4102.27	5.87
C18	Somalia	NULL	15.99
C19	UK	5087.38	66.78
C20	USA	10921.01	334.32
NULL	NULL	NULL	NULL




```
1 •   select global_health.Calculate_Avg_Healthcare_Expenditure('R1');
2
```

Result check R1 Countries :
Ethiopia – 75.11
Ghana – 193.22



What are the rates of child mortality for countries in South-East Asia (WHO)?

```sql
-- What are the child mortality rates and change
-- for countries in WHO_Region South-East Asia R3
SELECT
    country_id,
    country_name,
    Mortality2019_pc,
    Mortality1969_pc,
    (Mortality2019_pc - Mortality1969_pc) AS Child_Mortality_Change
FROM global_health.child_mortality cm
WHERE
    cm.country_id IN (SELECT
            rc.country_ID
        FROM
            Global_health.Representative_Countries rc
        Where
            rc.region_ID = 'R3')
ORDER BY Mortality2019_pc DESC;
```

Subquery used to filter and display from region R3 using Representative_countries table

Country_ID	Country_name	Region_ID
C01	Afghanistan	R5
C02	Argentina	R2
C03	Australia	R6
C04	Bangladesh	R3
C05	Brazil	R2
C06	Canada	R2
C07	China	R6
C08	Ethiopia	R1
C09	Finland	R4
C10	France	R4
C11	Germany	R4
C12	Ghana	R1
C13	Indonesia	R3
C14	India	R3
C15	Japan	R6
C16	Russia	R4
C17	Singapore	R6
C18	Somalia	R5
C19	UK	R4
C20	USA	R2
NULL	NULL	NULL

country_id	country_name	Mortality2019_pc	Mortality1969_pc	Child_Mortality_Change
C14	India	3.44	21.65	-18.21
C04	Bangladesh	3.07	22.87	-19.80
C13	Indonesia	2.38	17.19	-14.81

Qn: Based on the sample countries, which WHO region had the lowest life expectancy of both sexes combined?

Country_id	Country_name	REGION	Life_Expectancy_BothSexes
C08	Ethiopia	Africa	65.95
C12	Ghana	Africa	64.75
C02	Argentina	Americas	77.30
C05	Brazil	Americas	75.35
C06	Canada	Americas	82.35
C20	USA	Americas	79.15
C04	Bangladesh	South-East Asia	72.90
C13	Indonesia	South-East Asia	70.55
C14	India	South-East Asia	70.95
C09	Finland	Europe	81.85
C10	France	Europe	82.65
C11	Germany	Europe	81.55
C16	Russia	Europe	73.80
C19	UK	Europe	81.70
C01	Afghanistan	Eastern Medite...	63.65
C18	Somalia	Eastern Medite...	57.10
C03	Australia	Western Pacific	83.10
C07	China	Western Pacific	78.05
C15	Japan	Western Pacific	84.40
C17	Singapore	Western Pacific	83.75

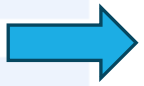
Average calculation between 2 column per row, formatted to 2 decimal places

```
SELECT
    c.Country_id,
    c.Country_name,
    r.REGION,
    FORMAT(AVG((l.Male_Life_Ex + l.Female_Life_Ex) / 2), 2) AS Life_Expectancy_BothSexes
FROM Life_Expectancy l
JOIN Representative_Countries c ON l.Country_ID = c.Country_ID
JOIN WHO_Regions r ON c.Region_ID = r.Region_ID
GROUP BY c.Country_id, c.Country_name, r.REGION;
```

Inner Join used as need to have “Region” for this query so any entries with missing “Regions” will not be included.

```sql
SELECT
    r.Region_ID,
    r.REGION,
    FORMAT(AVG((l.Male_Life_Ex + l.Female_Life_Ex) / 2), 2) AS Avg_Life_Expectancy_BothSexes
FROM WHO_Regions r
JOIN Representative_Countries c ON r.Region_ID = c.Region_ID
JOIN Life_Expectancy l ON c.Country_ID = l.Country_ID
GROUP BY r.Region_ID, r.REGION;
```

Country_id	Country_name	REGION	Life_Expectancy_BothSexes
C08	Ethiopia	Africa	65.95
C12	Ghana	Africa	64.75
C02	Argentina	Americas	77.30
C05	Brazil	Americas	75.35
C06	Canada	Americas	82.35
C20	USA	Americas	79.15
C04	Bangladesh	South-East Asia	72.90
C13	Indonesia	South-East Asia	70.55
C14	India	South-East Asia	70.95
C09	Finland	Europe	81.85
C10	France	Europe	82.65
C11	Germany	Europe	81.55
C16	Russia	Europe	73.80
C19	UK	Europe	81.70
C01	Afghanistan	Eastern Medite...	63.65
C18	Somalia	Eastern Medite...	57.10
C03	Australia	Western Pacific	83.10
C07	China	Western Pacific	78.05
C15	Japan	Western Pacific	84.40
C17	Singapore	Western Pacific	83.75



Result Grid | Filter Rows: | Export: | Wrap Cell

Region_ID	REGION	Avg_Life_Expectancy_BothSexes
R5	Eastern Mediterranean	60.38
R1	Africa	65.35
R3	South-East Asia	71.47
R2	Americas	78.54
R4	Europe	80.31
R6	Western Pacific	82.33

Creating a view of 3-4 base tables to provide a comprehensive health data summary for each cou

Country_ID	Country_name	Health_Expenditure	Male_Life_Ex	Female_Life_Ex	Mortality2019_pc
C02	Argentina	2198.88	73.90	80.70	0.91
C03	Australia	5294.46	81.10	85.10	0.37
C04	Bangladesh	123.29	70.70	75.10	3.07
C05	Brazil	1497.81	72.20	78.50	1.49
C06	Canada	5520.65	80.30	84.40	0.51
C07	China	880.19	75.30	80.80	0.79
C08	Ethiopia	75.11	63.10	68.80	5.08
C09	Finland	4710.00	79.20	84.50	0.23
C10	France	5492.53	79.70	85.60	0.43
C11	Germany	6738.67	79.10	84.00	0.37
C12	Ghana	193.22	62.70	66.80	4.64
C13	Indonesia	358.29	68.50	72.60	2.38
C14	India	211.00	69.50	72.40	3.44
C15	Japan	4587.03	81.40	87.40	0.25
C16	Russia	1704.04	68.80	78.80	0.58
C17	Singapore	4102.27	81.60	85.90	0.24
C18	Somalia	NULL	55.10	59.10	11.83
C19	UK	5087.38	79.90	83.50	0.43
C20	USA	10921.01	76.60	81.70	0.64

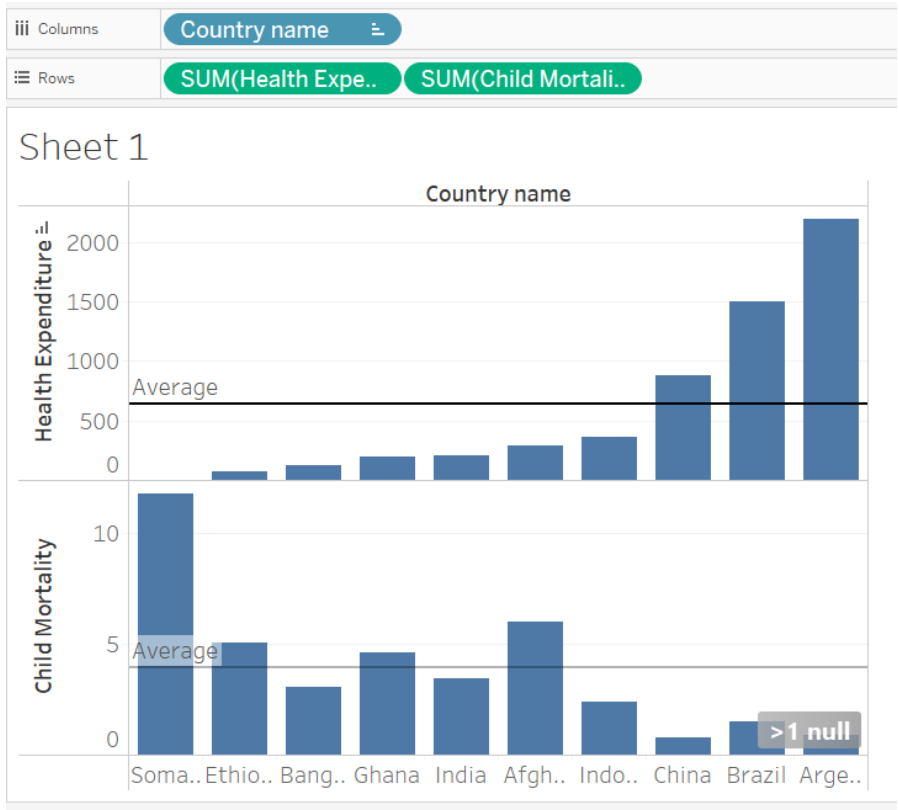
-
```sql
CREATE VIEW Health_Data_View AS
SELECT
    rc.Country_ID,
    rc.Country_name,
    ch.Health_Expenditure,
    le.Male_Life_Ex,
    le.Female_Life_Ex,
    cm.Mortality2019_pc
FROM Representative_Countries rc
LEFT JOIN Country_Healthex_Pop ch ON rc.Country_ID = ch.Country_ID
LEFT JOIN Life_Expectancy le ON rc.Country_ID = le.Country_Id
LEFT JOIN Child_Mortality cm ON rc.Country_ID = cm.Country_Id;
```

Find the top 10 countries with the highest child mortality rate and see if there is a possible correlation with health expenditure.

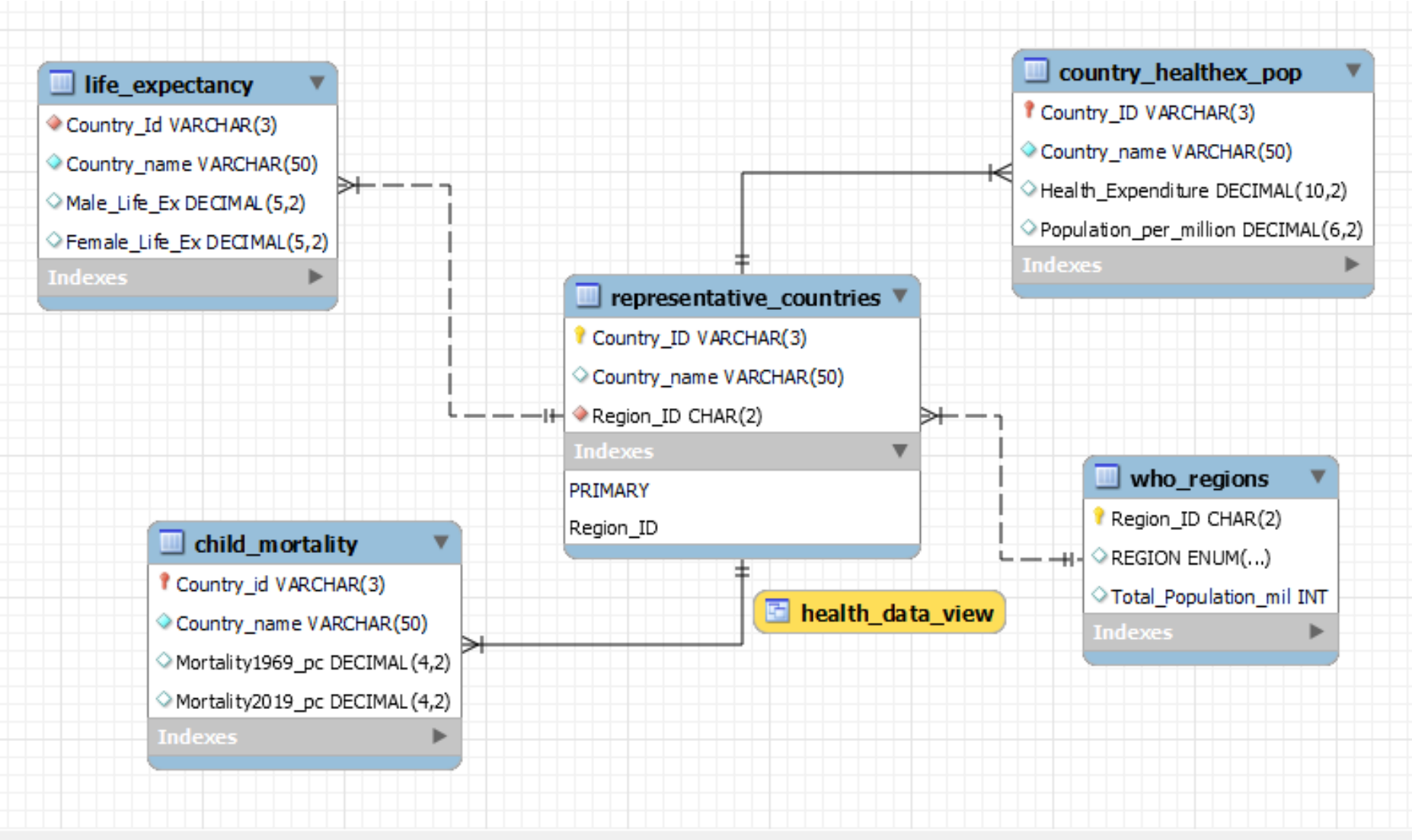
-
```sql
SELECT Country_name, Mortality2019_pc as Child_Mortality, Health_Expenditure
FROM Health_Data_View
ORDER BY Child_Mortality DESC
LIMIT 10;
```

Using Tableau to visualise data:

	Country_name	Child_Mortality	Health_Expenditure
	Somalia	11.83	NULL
▶	Afghanistan	6.01	285.56
	Ethiopia	5.08	75.11
	Ghana	4.64	193.22
	India	3.44	211.00
	Bangladesh	3.07	123.29
	Indonesia	2.38	358.29
	Brazil	1.49	1497.81
	Argentina	0.91	2198.88
	China	0.79	880.19



Child mortality is inversely co-related with health expenditure – countries which spend more on health show lower child mortality.


life_expectancy
Country_Id VARCHAR(3)
Country_name VARCHAR(50)
Male_Life_Ex DECIMAL(5,2)
Female_Life_Ex DECIMAL(5,2)
Indexes
country_healthex_pop
Country_ID VARCHAR(3)
Country_name VARCHAR(50)
Health_Expenditure DECIMAL(10,2)
Population_per_million DECIMAL(6,2)
Indexes
representative_countries
Country_ID VARCHAR(3)
Country_name VARCHAR(50)
Region_ID CHAR(2)
Indexes
PRIMARY
Region_ID
who_regions
Region_ID CHAR(2)
REGION ENUM(...)
Total_Population_mil INT
Indexes
child_mortality
Country_id VARCHAR(3)
Country_name VARCHAR(50)
Mortality1969_pc DECIMAL(4,2)
Mortality2019_pc DECIMAL(4,2)
Indexes
health_data_view

THANK YOU