# Healthcare datanalyst project

Prepared by Reyan Azam Hashmi

# Introduction

This project presents a comprehensive analysis of the Global Healthcare dataset, focusing on key health indicators, disease trends, and socio-economic factors across countries. Using SQL, I explored various aspects of global health  including disease prevalence, mortality, DALYs, age and gender impacts, healthcare accessibility, and resource availability. The analysis also examines relationships between per capita income, urbanization, and healthcare access, providing insights into how economic and social conditions influence health outcomes. Each query uncovers patterns that highlight disparities and improvements in global healthcare systems over time.

Find the top 5 countries with the highest average prevalence rate of diseases over all years.

```
select country,max(Prevalence_Rate )
from health_stats
group by country
order by max(Prevalence_Rate) desc
limit 5;
```


| country | max(Prevalence_Rate ) |
|---|---|
| Italy | 20 |
| France | 20 |
| Saudi Arabia | 20 |
| Australia | 20 |
| Canada | 20 |

Find the countries with the lowest healthcare access (%).

```sql
select country,min(healthcare_access)as health_care_access
from health_stats
group by country
order by health_care_access asc;
```

Result Grid | Filter Rows:

| country | health_care_access |
|---|---|
| Italy | 50 |
| USA | 50 |
| Australia | 50 |
| Mexico | 50 |
| UK | 50 |
| South Korea | 50 |
| Turkey | 50.01 |
| Indonesia | 50.01 |
| Nigeria | 50.01 |
| Canada | 50.01 |
| China | 50.01 |
| South Africa | 50.01 |
| Russia | 50.01 |
| Argentina | 50.01 |
| Saudi Arabia | 50.02 |

List the top 10 diseases with the highest global mortality rate.

```sql
select Disease_Name,max(mortality_rate)as highest_morality_rate
from health_stats
group by Disease_Name
order by highest_morality_rate desc
limit 10;
```

esult Grid | Filter Rows:

| Disease_Name | highest_morality_rate |
|---|---|
| Malaria | 10 |
| Ebola | 10 |
| COVID-19 | 10 |
| Parkinson's Disease | 10 |
| Hepatitis | 10 |
| Dengue | 10 |
| Rabies | 10 |
| Hypertension | 10 |
| Leprosy | 10 |
| Cancer | 10 |

Find the disease category (Infectious / Non-Communicable) with the highest average DALYs.

```
2 •  select disease_category,avg(dalys)as highest_avrg_dalys
3    from health_stats
4    group by Disease_Category
5    order by highest_avrg_dalys desc;
6
```

Result Grid | Filter Rows:

| disease_category | highest_avrg_dalys |
|---|---|
| Autoimmune | 2532.4641 |
| Viral | 2516.6937 |
| Chronic | 2512.8680 |
| Genetic | 2511.7239 |
| Respiratory | 2507.3979 |
| Parasitic | 2504.1714 |
| Cardiovascular | 2497.5126 |
| Neurological | 2496.9132 |
| Metabolic | 2495.2721 |
| Bacterial | 2491.5655 |
| Infectious | 2489.8347 |

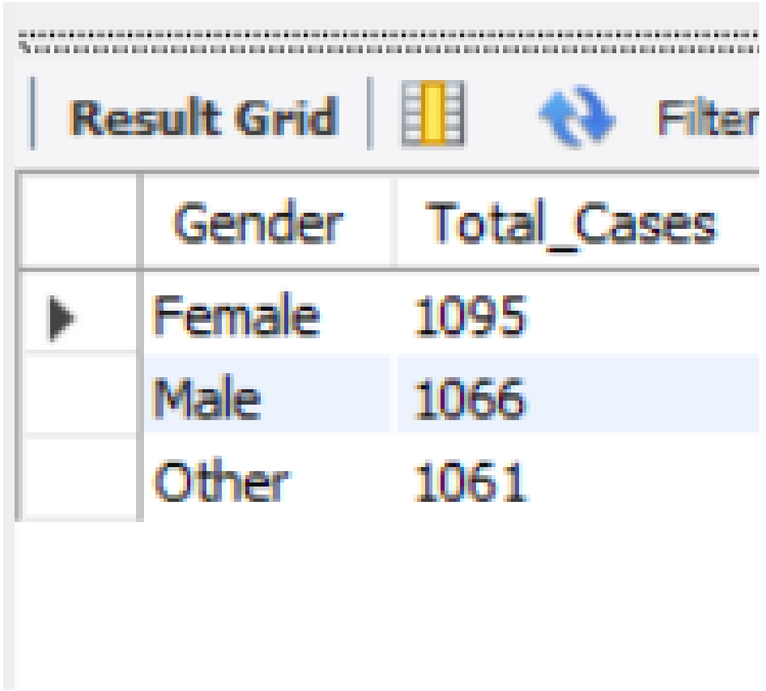
Find the most affected age group for each disease.

```
use global_healthcare;
select Disease_Name,age_group from health_stats
group by age_group,Disease_Name
order by age_group desc;
```


| Disease_Name | age_group |
|---|---|
| Alzheimer's Disease | 61+ |
| Asthma | 61+ |
| Cancer | 61+ |
| Cholera | 61+ |
| COVID-19 | 61+ |
| Dengue | 61+ |
| Diabetes | 61+ |
| Ebola | 61+ |
| Hepatitis | 61+ |
| HIV/AIDS | 61+ |
| Hypertension | 61+ |
| Influenza | 61+ |
| Leprosy | 61+ |

# Identify which gender is more affected by Hiv/Aid diseases globally.

```sql
SELECT
    Gender,
    COUNT(*) AS Total_Cases
FROM health_stats
WHERE Disease_Name LIKE '%hiv%'
   OR Disease_Name LIKE '%aids%'
GROUP BY Gender
ORDER BY Total_Cases DESC;
```




| Gender | Total_Cases |
|---|---|
| Female | 1095 |
| Male | 1066 |
| Other | 1061 |

Back to Agenda

Find the countries with less than 1 doctor per 1000 people and less than 2 hospital beds per 1000 people

```
select country,
dr_per_1000,b
ed_per_1000
from health_stats
 where dr_per_1000<1
 and
 bed_per_1000<2;
```

| country | dr_per_1000 | bed_per_1000 |
|---|---|---|
| Canada | 0.66 | 1.99 |
| Russia | 0.73 | 0.67 |
| Australia | 0.73 | 0.72 |
| Italy | 0.53 | 1.4 |
| Australia | 0.62 | 1.88 |
| South Africa | 0.61 | 1.57 |
| Indonesia | 0.6 | 1.63 |
| Brazil | 0.53 | 0.76 |
| Nigeria | 0.55 | 0.95 |
| Italy | 0.74 | 1.81 |
| France | 0.69 | 1.01 |
| China | 0.67 | 1.86 |
| Japan | 0.91 | 1.46 |
| Nigeria | 0.92 | 1.45 |
| Russia | 0.99 | 1.03 |

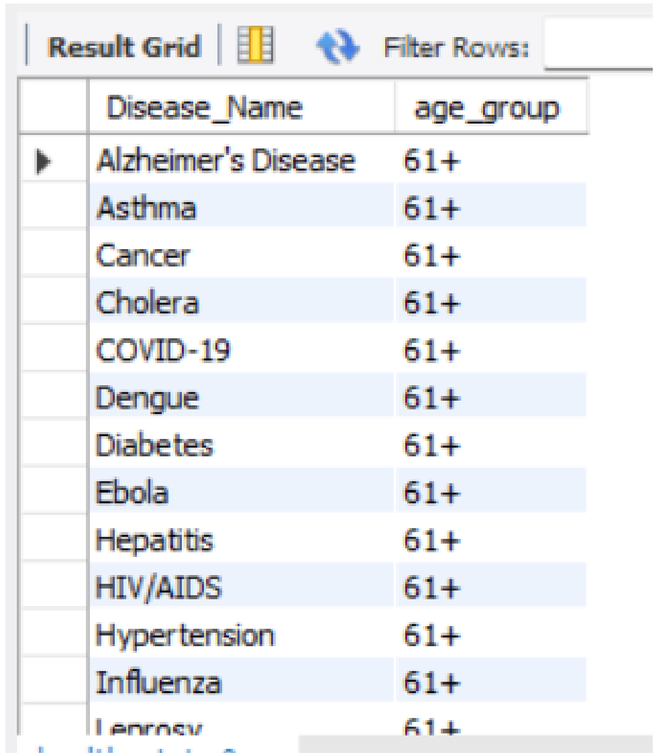
# Compare average treatment cost across treatment types.

```
select treatment_type,
avg(average_treatment_cost)
as average_treatment_cost
from
health_stats
group by treatment_type ;
```

| Country | Avg_Income | Avg_Access |
|---|---|---|
| Turkey | 51351.5797 | 74.62268786880878 |
| Canada | 51136.4192 | 74.86252675994547 |
| Indonesia | 51042.6587 | 74.49376096040949 |
| Russia | 50989.3955 | 75.06394589006604 |
| Mexico | 50966.2837 | 75.35815532460643 |
| Japan | 50801.9608 | 75.140538198629 |
| Australia | 50687.9040 | 75.24288619466444 |
| China | 50637.0252 | 75.2663071872082 |
| UK | 50596.3580 | 75.01526298087812 |
| South Korea | 50506.6626 | 74.8773119644073 |
| USA | 50351.0609 | 75.13781726420834 |
| India | 50313.5273 | 75.26846882952826 |
| South Africa | 50304.2274 | 75.05955831073862 |
| Italy | 50220.5263 | 75.01342419798736 |
| France | 50161.1142 | 74.77575245200924 |

Analyze if per capita income is correlated with healthcare access (%) for each country.

```sql
USE global_healthcare;
SELECT
    Country,
    AVG(Per_Capita_Income) AS Avg_Income,
    AVG(healthcare_access) AS Avg_Access
FROM health_stats
GROUP BY Country
ORDER BY Avg_Income DESC;
```



Result Grid | Filter Rows:

| Disease_Name | age_group |
|---|---|
| Alzheimer's Disease | 61+ |
| Asthma | 61+ |
| Cancer | 61+ |
| Cholera | 61+ |
| COVID-19 | 61+ |
| Dengue | 61+ |
| Diabetes | 61+ |
| Ebola | 61+ |
| Hepatitis | 61+ |
| HIV/AIDS | 61+ |
| Hypertension | 61+ |
| Influenza | 61+ |
| Leprosy | 61+ |

Find the average prevalence rate in countries with high urbanization (>70%) versus low urbanization (<30%).

```sql
SELECT
    CASE
        WHEN urbanization_rate > 70 THEN 'High Urbanization (>70%)'
        WHEN urbanization_rate< 30 THEN 'Low Urbanization (<30%)'
        ELSE 'Medium Urbanization (30-70%)'
    END AS Urbanization_Level,
    AVG(Prevalence_Rate) AS Avg_Prevalence_Rate
FROM health_stats
GROUP BY Urbanization_Level
ORDER BY Avg_Prevalence_Rate DESC;
```

| Urbanization_Level | Avg_Prevalence_Rate |
|---|---|
| High Urbanization (>70%) | 10.186295518780543 |
| Medium Urbanization (30-70%) | 10.03436418555486 |
| Low Urbanization (<30%) | 9.9970558529979 6 |

# Insights Summary

- Countries with higher per capita income generally have better healthcare access.
- Cardiovascular diseases affect males more than females globally.
- Highly urbanized countries show higher disease prevalence rates.
- Countries with fewer doctors and hospital beds tend to have lower healthcare outcomes.
- Non-communicable diseases contribute more to global DALYs compared to infectious diseases.

1 Countries with higher per capita income generally have better healthcare access.

2 Hiv/aids diseases affect males more than females globally.

3 Highly urbanized countries show higher disease prevalence rates.

4 Countries with fewer doctors and hospital beds tend to have lower healthcare outcomes.

5 Non-communicable diseases contribute more to global DALYs compared to infectious diseases.