

Predictive Analytics MGS 616 Group Project

Team FP_8: Sai Kartheek Mahankali, Nihareeka Suneel Gote, Viraj Vhatkar, Aditya Manjunath Naik

Datasets

1. [WHO Mortality Data – All causes](#)
2. [WHO Mortality Data – Communicable Diseases](#)
3. [WHO Mortality Data – Non-Communicable Diseases](#)
4. [Annual number of deaths by cause](#)
5. [Global Health Expenditure Data](#)

Data Description

1. WHO Mortality Data – All causes:

The dataset is sourced from [WHO](#) and consists of 300K rows and 13 columns namely - Region Code, Region Name, Country Code, Country Name, Year, Sex, Cause of Death, Age group code, Age Group, Number, Percentage of cause-specific deaths out of total deaths, Age-standardized death rate per 100 000 standard population, Death rate per 100 000 population.

The data is in CSV format and has details in the time period 1950 – 2020.

2. WHO Mortality Data – Communicable Diseases:

The dataset is sourced from [WHO](#) and consists of 298K rows and 13 columns namely - Region Code, Region Name, Country Code, Country Name, Year, Sex, Cause of Death, Age group code, Age Group, Number, Percentage of cause-specific deaths out of total deaths, Age-standardized death rate per 100 000 standard population, Death rate per 100 000 population.

The data is in CSV format and has details in the time period 1950 – 2020.

3. WHO Mortality Data – Non-Communicable Diseases:

The dataset is sourced from [WHO](#) and consists of 298K rows and 13 columns namely - Region Code, Region Name, Country Code, Country Name, Year, Sex, Cause of Death, Age group code, Age Group, Number, Percentage of cause-specific deaths out of total deaths, Age-standardized death rate per 100 000 standard population, Death rate per 100 000 population.

The data is in CSV format and has details in the time period 1950 – 2020.

4. Annual number of deaths by cause:

The dataset is sourced from [ourworldindata.org](#) and consists of 6K rows and 36 columns namely – Entity, Code, Year, Number of executions (Amnesty International), Deaths - Meningitis, Deaths - Alzheimer's disease and other dementias, Deaths - Parkinson's disease, Deaths - Nutritional deficiencies, Deaths - Malaria, Deaths - Drowning, Deaths - Interpersonal violence, Deaths - Maternal disorders, Deaths - HIV/AIDS, Deaths - Drug use disorders, Deaths - Tuberculosis, Deaths - Cardiovascular diseases, Deaths - Lower respiratory infections, Deaths - Neonatal disorders, Deaths - Alcohol use disorders, Deaths - Self-harm, Deaths - Exposure to forces of nature, Deaths - Diarrheal diseases, Deaths - Environmental heat and cold exposure, Deaths - Neoplasms, Deaths - Conflict and terrorism, Deaths - Diabetes mellitus, Deaths - Chronic kidney disease, Deaths - Poisonings, Deaths - Protein-energy malnutrition, Deaths - Terrorism, Deaths -

Road injuries, Deaths - Chronic respiratory diseases, Deaths - Cirrhosis and other chronic liver diseases, Deaths - Digestive diseases, Deaths - Fire, heat, and hot substances, Deaths - Acute hepatitis.

The data is in CSV format and has details in the time period 1990 – 2019

5. Global Health Expenditure Data:

The dataset is sourced from [WHO](#) and consists of 4K rows and mainly 13 columns as below -

- che_pc_usd - Current Health Expenditure (CHE) per Capita in US\$
- che_usd - Current Health Expenditure (CHE), in million current US\$
- gghed_usd - Domestic General Government Health Expenditure (GGHE-D), in million current US\$
- pvtd_usd - Domestic Private Health Expenditure (PVT-D), in million current US\$
- ext_usd - External Health Expenditure (EXT), in million current US\$
- che_usd2020 - Current Health Expenditure (CHE), in million constant (2020) US\$
- gghed_usd2020 - Domestic General Government Health Expenditure (GGHE-D), in million constant (2020) US\$
- pvtd_usd2020 - Domestic Private Health Expenditure (PVT-D), in million constant (2020) US\$
- ext_usd2020 - External Health Expenditure (EXT), in million constant (2020) US\$
- che_usd2020_pc - Current Health Expenditure (CHE), in constant (2020) US\$ per capita
- gghed_usd2020_pc - Domestic General Government Health Expenditure (GGHE-D), in constant (2020) US\$ per capita
- pvtd_usd2020_pc - Domestic Private Health Expenditure (PVT-D), in constant (2020) US\$ per capita
- ext_usd2020_pc - External Health Expenditure (EXT), in constant (2020) US\$ per capita

The data is in Excel (.xlsx) format and has details in the time period 2000 – 2021