# Information Architecture Final Project

# under-5 child mortality

## Group 4 - Team Members

* Mazal-tov Amsellen
* Manling Yang
* Julian Ruggiero

## Context

Among the most fundamental children rights is the right to survive. While substantial progress in child survival has been made in the past 30 years, in 2018 alone, estimated 5.3 million children died before reaching their fifth birthday. We aim to understand how this problematic is impacting different regions and countries around the world along with the main reasons of death.

## Architecture in AWS

## Data sources:

### 1. World Development Indicators (WDI)

The data we get from WDI API is related to:

* Regions (Latin America, South Asia, Sub-Saharan Africa, etc.).
* Countries.
* Income Groups per year (High income, Low income, Lower middle income, Upper middle income).
* Indicator 1: Under 5 child mortality rates per 1000 births by year and by country.
* Indicator 2: Gross National Product (GNI) per year per country.

Once these entities are gathered, we save csv files to s3 folder 'WDI' running a Lambda function, main code is in Scripts/WDI\_API\_to\_S3.py (To explore the WDI we use world\_bank\_data library, the API v2 implementation in python.)

<https://pypi.org/project/world-bank-data/>

### 2. World Health Organization (WHO)

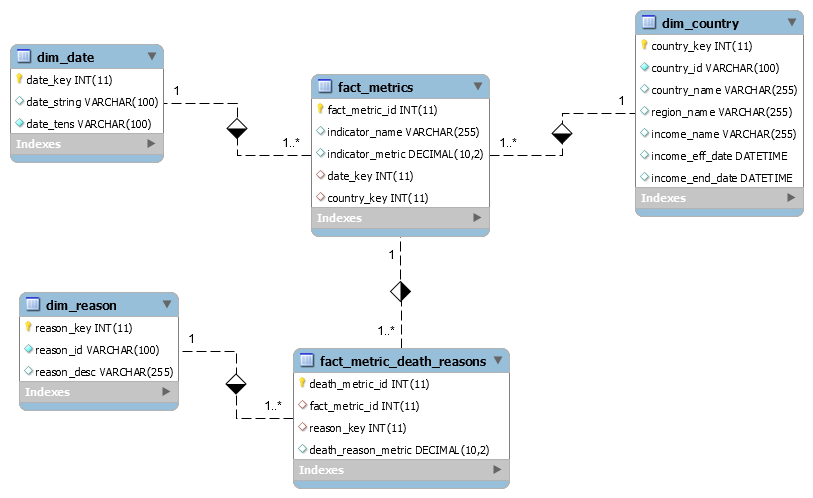
We access the Global Health Observatory data repository under the URL below: <https://apps.who.int/gho/data/node.main.ChildMortCTRY2002015?lang=en>

In this website we can find Child mortality's Rate of deaths by cause and by country, starting with HIV/AIDS.

We use web scraping with Python and Selenium to download the corresponding information for each of the 12 remaining reasons below, accessing each element in the list and taking the 'href' value to navigate into the next page:

* Diarrhoeal diseases
* Measles
* Meningitis/encephalitis
* Malaria
* Acute lower respiratory infections
* Prematurity
* Birth asphyxia and birth trauma
* Sepsis and other infectious conditions of the newborn
* Other communicable, perinatal and nutritional conditions
* Congenital anomalies
* Other noncommunicable diseases

## Database design – star schema



## Visualizations

Main dashboard can be found in the link below

<https://public.tableau.com/profile/julian.ruggiero#!/vizhome/ChildMortality_15926907418080/Dashboardjuli?publish=yes>