

# Data Description

We used data from the [World Bank's World Development Indicators \(WDI\)](#) to examine population dynamics across five countries: Afghanistan, Kenya, India, Brazil, and the United Kingdom. These countries were selected to represent income levels ranging from low-income to high-income, based on the [World Bank's Classification of the world by income and region](#). Specifically, Afghanistan is low-income, Kenya and India are lower-middle income, Brazil is upper-middle income, and the United Kingdom is classified as high-income. This allows for meaningful comparisons across development contexts.

We focused specifically on three indicators ranging from 2000 to 2023. Below are the indicators and their unique WDI codes:

1. Life expectancy at birth (SP.DYN.LE00.IN)
2. Adolescent fertility rate (SP.ADO.TFRT)
3. Under-5 mortality rate (SH.DYN.MORT)

We also included the secondary school enrollment indicator (SE.SEC.ENRR) to use for our analysis on adolescent fertility.

The data was accessed via the `wbdata` Python package. First, we used Python to import the necessary packages and select our countries and indicators. Then, we extracted the year from the date column, and filtered the data to include only observations from the selected countries and relevant years.

We used the `duckdb` package to create a view (`wdi_raw`) and query and transform data using SQL syntax within Python. Using SQL syntax, we selected the country, year, and indicator columns. We took the maximum of each indicator to avoid duplicates, and grouped and ordered by country and year to organize our data. We then returned this dataset as a Pandas dataframe before saving it as a csv file to our data folder.

The final variables in our dataset are:

- country: country name (e.g., "India")
- year: year of observation (2000–2023)
- life\_expectancy: average years at birth
- adolescent\_fertility: births per 1,000 women aged 15–19

- under5\_mortality: deaths per 1,000 live births
- secondary\_school\_enrollment: ratio of students enrolled in secondary school to the population of the age group

Below are our summary statistics for our three primary indicators:

### Average Life Expectancy by Country (2000–2023)

Country	Mean	Min	Median	Max
United Kingdom	80.09	77.74	80.52	81.37
Brazil	73.44	69.58	73.91	75.85
India	67.45	62.75	67.45	72.00
Kenya	60.53	56.08	61.25	63.65
Afghanistan	60.49	55.00	60.98	66.04

### Under-5 Mortality Rate by Country (2000–2023)

Country	Mean	Min	Median	Max
United Kingdom	5.18	4.40	4.90	6.50
Brazil	20.18	14.40	17.55	34.50
India	55.70	27.70	53.50	91.60
Kenya	58.16	39.90	51.70	96.30
Afghanistan	87.44	55.50	82.90	131.70

### Adolescent Fertility Rate by Country (2000–2023)

Country	Mean	Min	Median	Max
United Kingdom	19.28	8.36	20.88	28.73
India	46.11	14.06	40.16	114.35
Brazil	64.43	42.69	64.53	85.61
Kenya	93.82	56.31	102.42	123.25
Afghanistan	102.32	64.07	98.45	154.31