[redacted]

```
from google.colab import files
uploaded = files.upload()
```

Choose Files No file chosen   Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

```
Saving unicef_indicator_1.csv to unicef_indicator_1.csv
```

```
import pandas as pd

# Read the uploaded files
indicator_df = pd.read_csv('unicef_indicator_1.csv')
metadata_df = pd.read_csv('unicef_metadata.csv')

# Preview the first few rows of both files
print("Indicator data:")
print(indicator_df.head())

print("\nMetadata:")
print(metadata_df.head())
```

```
Indicator data:
          country alpha_2_code alpha_3_code numeric_code\
0 Afghanistan            AF          AFG             4
1 Afghanistan            AF          AFG             4
2 Afghanistan 3          AF          AFG              4
Afghanistan   4          AF          AFG              4
Afghanistan              AF          AFG             4

                                    indicator time_period obs_value  \
0 Diarrhoea treatment - percentage of children (...          2011        51.3
1 Diarrhoea treatment - percentage of children (...          2015        46.4
2 Diarrhoea treatment - percentage of children (...          2018        41.9
3 Diarrhoea treatment - percentage of children (...          2023        31.9
4 Diarrhoea treatment - percentage of children (...          2011        55.1

      sex unit_multiplier unit_of_measure observation_status \
 0 Female          Units               %           Reported
 1 Female          Units               %           Reported
 2 Female          Units               %           Reported
 3 Female          Units               %           Reported
4    Male          Units               %           Reported

   observation_confidentaility  \
0                         Free
1                         Free
2                         Free
3                         Free
```

```
4                        Free

  time_period_activity_related_to_when_the_data_are_collected  \
0                                   End of fieldwork
1                                   End of fieldwork
2                                   End of fieldwork
3                                   End of fieldwork
4                                   End of fieldwork

          current_age
  0 Under 5 years old
  1 Under 5 years old
  2 Under 5 years old
  3 Under 5 years old
  4 Under 5 years old

Metadata:
         country alpha_2_code alpha_3_code numeric_code year  \
0 Afghanistan           AF          AFG             4 1960
1 Afghanistan           AF          AFG             4 1961
2 Afghanistan           AF          AFG             4 1962
3 Afghanistan           AF          AFG             4 1963
4 Afghanistan           AF          AFG             4 1964

   Population, total  GDP per capita (constant 2015 US$) GNI (current US$) \
0          9035043.0                                 NaN      5.488888e+08
1          9214083.0                                 NaN      5.600000e+08
2          9404406.0                                 NaN      5.577778e+08
3          9604487.0                                 NaN      7.666667e+08
4          9814318.0                                 NaN      8.155556e+08
```

```
!pip install plotnine

# Import necessary libraries
from plotnine import *
```

```
Requirement already satisfied: plotnine in /usr/local/lib/python3.11/dist-packages (0.14.5)
Requirement already satisfied: matplotlib>=3.8.0 in /usr/local/lib/python3.11/dist-packages (from plotnine) (3.10.0)
Requirement already satisfied: pandas>=2.2.0 in /usr/local/lib/python3.11/dist-packages (from plotnine) (2.2.2)
Requirement already satisfied: mizani~=0.13.0 in /usr/local/lib/python3.11/dist-packages (from plotnine) (0.13.3)
Requirement already satisfied: numpy>=1.23.5 in /usr/local/lib/python3.11/dist-packages (from plotnine) (2.0.2)
Requirement already satisfied: scipy>=1.8.0 in /usr/local/lib/python3.11/dist-packages (from plotnine) (1.14.1)
Requirement already satisfied: statsmodels>=0.14.0 in /usr/local/lib/python3.11/dist-packages (from plotnine) (0.14.4)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.11/dist-packages (from matplotlib>=3.8.0->plotnine) (1.3.2)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.11/dist-packages (from matplotlib>=3.8.0->plotnine) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.11/dist-packages (from matplotlib>=3.8.0->plotnine) (4.57.0)
Requirement already satisfied: kiwisolver>=1.3.1 in /usr/local/lib/python3.11/dist-packages (from matplotlib>=3.8.0->plotnine) (1.4.8)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-packages (from matplotlib>=3.8.0->plotnine) (24.2)
Requirement already satisfied: pillow>=8 in /usr/local/lib/python3.11/dist-packages (from matplotlib>=3.8.0->plotnine) (11.1.0)
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.11/dist-packages (from matplotlib>=3.8.0->plotnine) (3.2.3)
Requirement already satisfied: python-dateutil>=2.7 in /usr/local/lib/python3.11/dist-packages (from matplotlib>=3.8.0->plotnine) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-packages (from pandas>=2.2.0->plotnine) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-packages (from pandas>=2.2.0->plotnine) (2025.2)
Requirement already satisfied: patsy>=0.5.6 in /usr/local/lib/python3.11/dist-packages (from statsmodels>=0.14.0->plotnine) (1.0.1)
```

```
     Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.11/dist-packages (from python-dateutil>=2.7->matplotlib>=3.8.0->plotnine) (1.17.0)
```

```
from google.colab import files

# Upload your image (select from your computer)
uploaded = files.upload()
```

Choose Files No file chosen Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

□ *UNICEF Data Analysis Report*

Raising Awareness Through Global Health Indicators

```
from IPython.display import Image, display

# Show your uploaded image (with the correct filename)
display(Image(filename='diarrhoea treatment image.png'))
```

Introduction

Every child deserves access to basic healthcare, yet millions around the world still lack it. One of the clearest examples of this is the treatment of diarrhoeal diseases—a leading cause of death among children under five.

This report, developed for UNICEF, explores global disparities in diarrhoea treatment, and links them with socioeconomic indicators such as GDP per capita and life expectancy. By analyzing public UNICEF datasets, we aim to raise awareness about where the gaps exist, and what

global development data can tell us about a child’s access to care.

Top 10 Countries – Diarrhoea Treatment in Girls (2023)

In 2023, significant progress was made in some countries toward providing diarrhoea treatment to young girls. However, the data shows stark inequalities. The bar chart below highlights the top 10 countries leading in female child treatment rates, revealing potential models for global improvement.

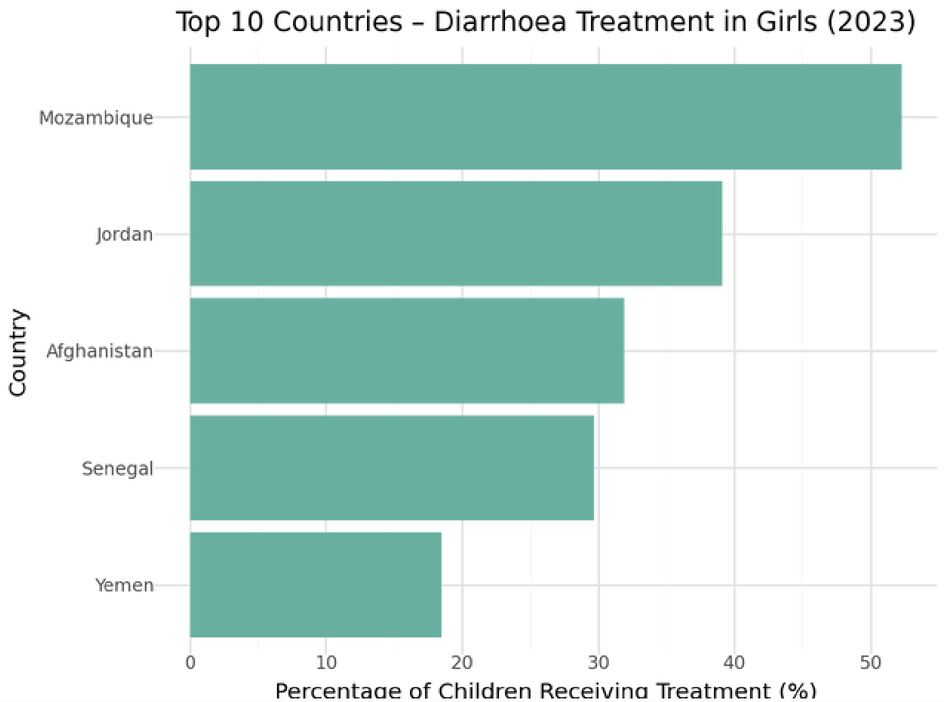
```
# Filter the dataset for Diarrhoea treatment data for female children in 2023
filtered = indicator_df[
    (indicator_df['indicator'].str.contains('Diarrhoea')) &
    (indicator_df['time_period'] == 2023) &
    (indicator_df['sex'] == 'Female')
]

# Get top 10 countries with highest treatment percentage
top10 = filtered.sort_values(by='obs_value', ascending=False).head(10)

# Create bar chart using plotnine
from plotnine import *

bar_chart = (
    ggplot(top10, aes(x='reorder(country, obs_value)', y='obs_value')) +
    geom_col(fill='#69b3a2') +
    coord_flip() +
    labs(
        title='Top 10 Countries – Diarrhoea Treatment in Girls (2023)',
        x='Country',
        y='Percentage of Children Receiving Treatment (%)'
    ) +
    theme_minimal()
)

# Display the chart
bar_chart
```



☐ *Life Expectancy vs GDP per Capita*

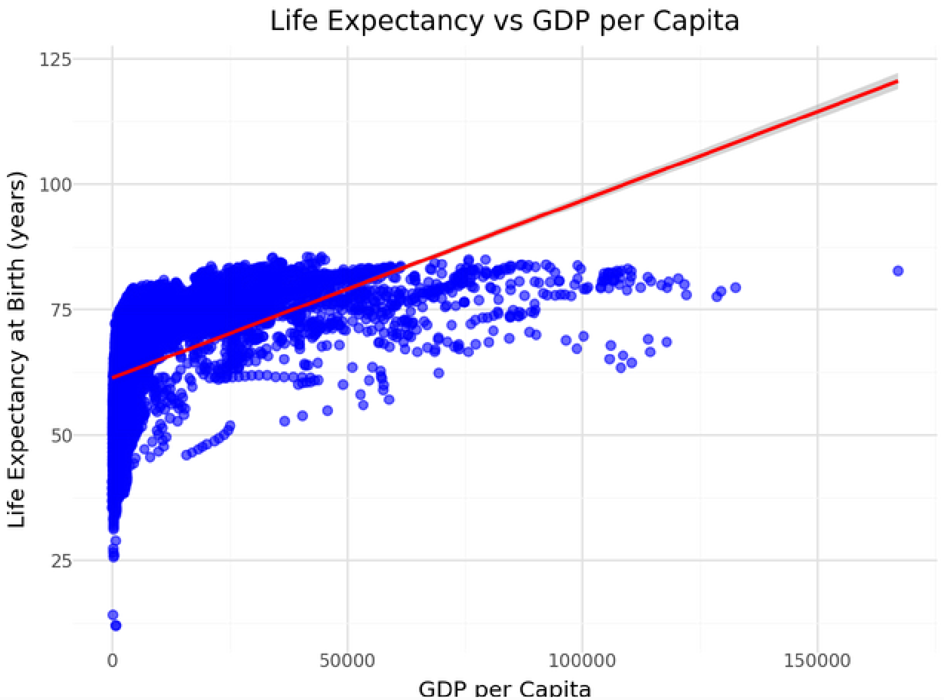
Economic strength is often linked with longer lives. The scatterplot shows a clear correlation between GDP per capita and life expectancy suggesting that wealthier nations tend to provide better healthcare and living conditions. The regression line emphasizes this relationship, raising questions about what can be done to support lower-income countries.

```
# Make a copy and rename column for clarity
scatter_df = metadata_df[
    ['country', 'GDP per capita (constant 2015 US$)', 'Life expectancy at birth, total (years)']
].dropna().rename(columns={
    'GDP per capita (constant 2015 US$)': 'GDP per capita'
})

# Import plotnine
from plotnine import *

# Create the scatterplot
scatter_plot = (
    ggplot(scatter_df,  aes(  x='GDP  per  capita',
    y='Life expectancy at birth, total (years)'
    )) +
    geom_point(color='blue', size=2, alpha=0.6) +
    geom_smooth(method='lm', color='red') +
    labs(
        title = "Life Expectancy vs GDP per Capita",
        x = "GDP per Capita",
        y = "Life Expectancy at Birth (years)"
    ) +
    theme_minimal()
)

# Show the chart
scatter_plot
```



□ *Diarrhoea Treatment Over Time (2011–2023)*

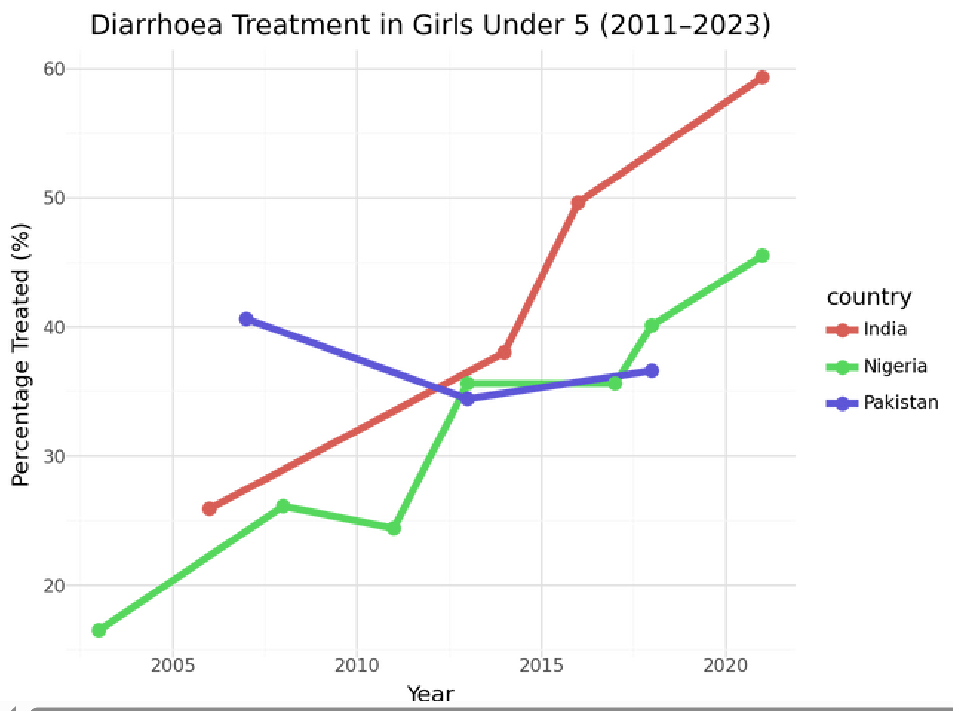
While some countries show gradual improvement in healthcare access, others face stagnation or decline. The time-series chart below highlights these trends in three populous countries—India, Nigeria, and Pakistan; showing that even within developing nations, progress is not always steady.

```
# Filter for diarrhoea treatment in girls under 5 in India, Nigeria, and Pakistan
countries = ['India', 'Nigeria', 'Pakistan']

time_df = indicator_df[
    (indicator_df['indicator'].str.contains('Diarrhoea')) &
    (indicator_df['sex'] == 'Female') &
    (indicator_df['country'].isin(countries))
]

# Create a time-series line chart
time_chart = (
    ggplot(time_df, aes(x='time_period', y='obs_value', color='country')) +
    geom_line(size=2) +
    geom_point(size=3) +
    labs(
        title='Diarrhoea Treatment in Girls Under 5 (2011–2023)',
        x='Year',
        y='Percentage Treated (%)'
    ) +
    theme_minimal()
)

# Show the chart
time_chart
```



Global Map of Diarrhoea Treatment Rates (Female, 2023)

Mapping the data makes the inequality stark. Some regions, particularly in Sub-Saharan Africa and parts of South Asia, continue to report low treatment rates. This map visualization draws attention to where international efforts must focus in improving basic child healthcare access.

```
# Install geopandas
!pip install geopandas

import geopandas as gpd
import pandas as pd
import matplotlib.pyplot as plt

# Load world map from a public GeoJSON
world = gpd.read_file("https://raw.githubusercontent.com/johan/world.geo.json/master/countries.geo.json")

# Filter treatment data for females in 2023
map_data = indicator_df[
    (indicator_df['indicator'].str.contains('Diarrhoea')) &
    (indicator_df['sex'] == 'Female') &
    (indicator_df['time_period'] == 2023)
]
```

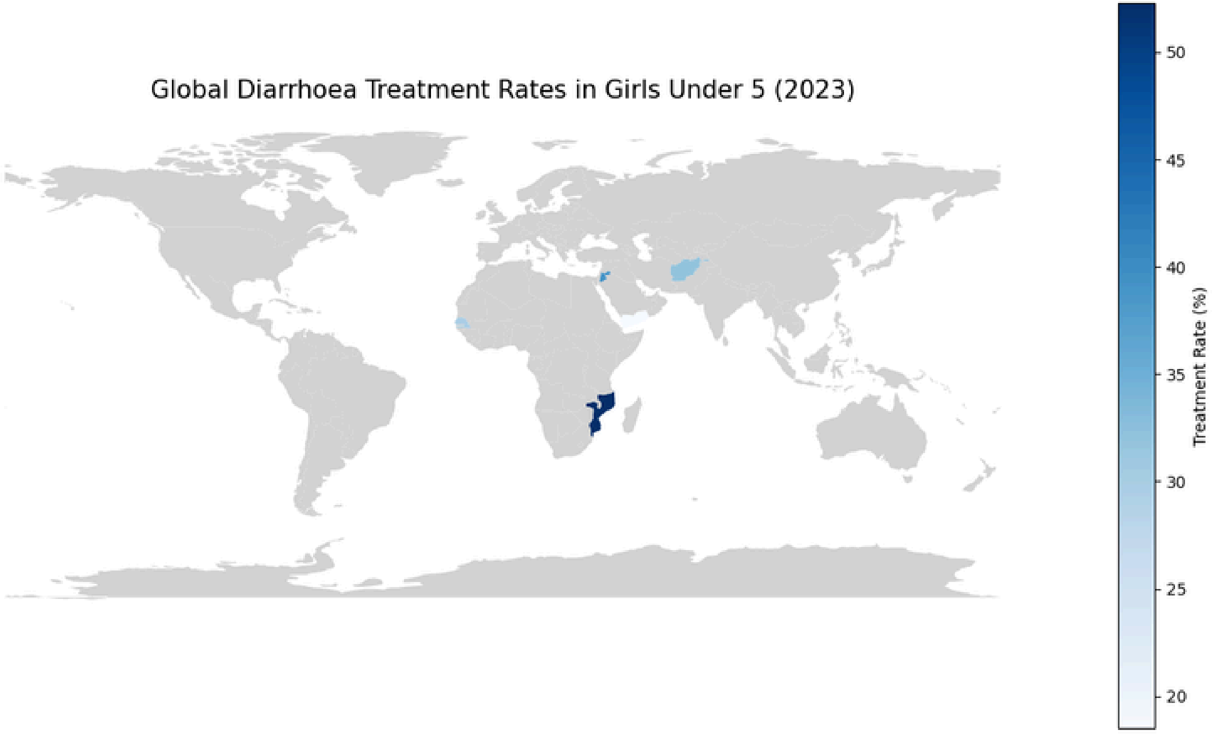
```
# Average value per country
map_data = map_data.groupby('country', as_index=False)['obs_value'].mean()

# Merge with GeoDataFrame
world = world.rename(columns={"name": "country"})
merged = world.merge(map_data, on="country", how="left")

# Plot world map
fig, ax = plt.subplots(1, 1, figsize=(15, 8))
merged.plot(column='obs_value', ax=ax, legend=True, cmap='Blues',
            legend_kwds={'label': "Treatment Rate (%)"},
            missing_kwds={"color": "lightgrey", "label": "No data"})

# Title and clean plot
ax.set_title("Global Diarrhoea Treatment Rates in Girls Under 5 (2023)", fontsize=15)
ax.axis("off")
plt.show()
```

```
Requirement already satisfied: geopandas in /usr/local/lib/python3.11/dist-packages (1.0.1)
Requirement already satisfied: numpy>=1.22 in /usr/local/lib/python3.11/dist-packages (from geopandas) (2.0.2)
```


Global Diarrhoea Treatment Rates in Girls Under 5 (2023)
50
45
40
35
30
25
20
Treatment Rate (%)