Number of HIV related deaths

(2000- 2021)

Presented by: Sahana Muralidaran (21076516) – MSc. Data Science (SW)

# Introduction

This report contains various data visualisation techniques related to the number of HIV related deaths that happened all over the world between the year 2000 and 2021.

The data for this report was taken from the WHO (World Health Organization) Global Health Observatory. The data is available under the “Number of People Dying from HIV- related causes”, as a CSV file. The web link for the data is below:

<https://www.who.int/data/gho/data/indicators/indicator-details/GHO/number-of-deaths-due-to-hiv-aids>

The data has a total of 4268 rows and 34 columns, out of which 1280 rows were none values. So, eliminating those rows we have a total of 2988 rows of data. Similarly, we can eliminate columns like ('IndicatorCode','Indicator', 'ValueType','Location type','Period type', 'Dim1 type', 'Dim1', 'Dim1ValueCode', 'Dim2 type','Dim2','Dim2ValueCode', 'Dim3 type',' Dim3', 'Dim3ValueCode', 'DataSourceDimValueCode','DataSource', 'FactValueNumericPrefix', 'Value', 'FactValueUoM', 'FactValueNumericLowPrefix', 'FactValueNumericHighPrefix', 'FactValueTranslationID', 'FactComments', 'Language','DateModified', 'IsLatestYear'), which are either empty or of less significance.

So, after cleaning the data we are left with data of number of HIV related deaths for 135 countries, which spans for 2988 rows and 8 columns.

The data available is classified as 6 regions, those are:

* Africa
* Americas
* Eastern Mediterranean
* Western Pacific
* Europe
* South-East Asia

# Methodology

We can use various visualization techniques to understand the global trend for the number of HIV related deaths between 2000 and 2021.

## Line plot

This is a line plot showing the number of HIV deaths for every year between 2000 and 2021 for all the 6 regions. The X axis has years (2000-2021) and the Y axis represents the number of deaths.

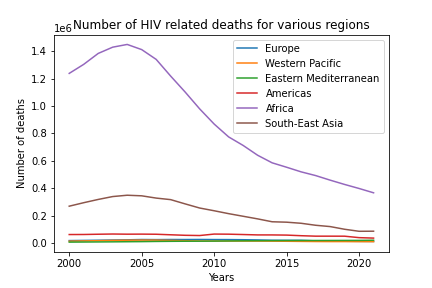


Figure 1

From this graph we can observe that:

* The number of deaths in Africa has always been higher than the rest of the world.
* The scale of the y axis has been altered according that of Africa, therefore the line representing Western Pacific, Eastern Mediterranean, Americas and Europe are close to zero. Since the number of deaths in these regions are a few thousand compared to that of Africa where the deaths are several hundreds of thousands.

## Subplot

The below figure shows six subplots (one for each region). The X axis has years (2000-2021) and the Y axis represents the number of deaths.

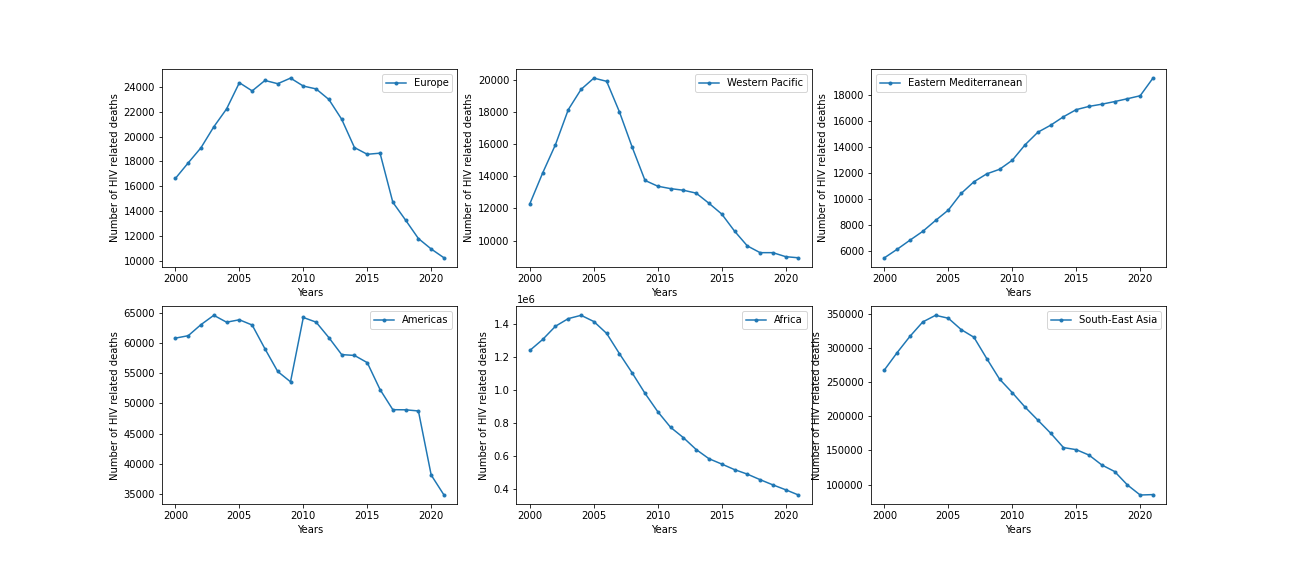


Figure 2

From this graph we can observe that:

### Europe

* The graph shows a steep increase from 2000 to 2005, where it went from close to 16000 to 24000.
* There was a slow increase from 2005 to 2010 and it reached its all-time maximum.
* After 2010, there was a drastic drop in the numbers and it reached close to 10000 in the year 2021.
* The number of deaths was almost equal for 2015 and 2016. (Around 18000)

### Western Pacific

* The graph shows a steep increase from around 12000 to 20000 between the years 2000 to 2005.
* From 2005 to 2009, there was a steep drop in graph (number went from 20000 to 14000
* After 2009, there was a steady drop in numbers, and the total count went below 10000.

### Americas

* The number of deaths ranges between 60000 to 65000 from 2000 to 2006.
* The number dropped by almost 10000 in the next 2 to 3 years from 2006.
* After the drop, there was an equal amount (almost 10000) of increase in the deaths till 2010.
* For the next 6 years, the number went from 65000 to 50000 and becomes stable till 2019.
* Within the next two years (2019-2021), there was a sudden drop and the number went below 35000

### Africa

* The graph was smooth overall.
* The graph reached a maximum of more than 1.4 million in the year 2004, after that there was a steady drop in numbers till 2021.
* Interestingly, the number of deaths decreased by more than a million in 16 years.

### South-East Asia

* The graph for this region is similar to that of Africa.
* The graph reached a maximum of close to 350000 in the year 2004 and then dropped steadily till 2021 and became less than 100000. (i.e., a drop of 250000 within 16 years)

### Eastern Mediterranean

* Unlike the other graphs, this graph shows a linear increase from 2000 to 2021.
* The numbers went from less than 6000 in the year 2000 to more than 18000 in 2021.
* It had an increase of more than 12000 within 16 years, which is approximately an increase of 750 per year.

## Pie chart

The below figure shows the top 8 countries with the highest number of HIV related deaths between the years 2000 and 2021.

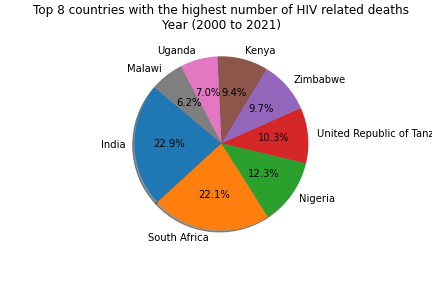


Figure 3

From this plot we can observe that:

* India has the highest number of deaths followed by South Africa, Nigeria, United republic of Tanzania, Zimbabwe, Kenya, Uganda, Malawi.
* Though we found that the region Africa had the highest number of deaths, we can observe that India, which belongs to the South-East Asia region has the highest number of overall deaths.
* Countries from Africa region occupy the remaining 7 positions in the ranking shown above.
* We can also note that the number of deaths in India and South Africa are almost similar.

## Bar Plot

The below figure shows the bar plot for the number of HIV related deaths across the world. The X axis has years (2000-2021) and the Y axis represents the number of deaths

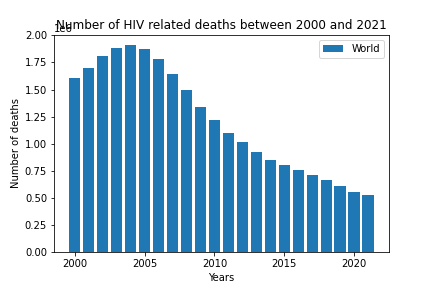


Figure 4

From this graph we can observe that:

* The graph was smooth increase from 2000 to 2004, where the numbers increased by close to 0.5 million.
* The graph reached a maximum of more than 1.8 million in the year 2004, after that there was a steady drop in numbers till 2021.
* Interestingly, the number of deaths decreased by close to 1.5 million in 16 years starting from 2004 to 2021.

# Summary

From the above analysis we can conclude:

* There was an overall maximum in the year 2004 after which the numbers have reduced steadily.
* The decrease in number might be due to the various initiatives taken for the prevention of HIV like the “access to all’ theme introduced by the International AIDS Conference and 3-by-5 initiative by WHO.
* Africa has always had the highest number of deaths and it reached a maximum value of more than 1.4 million in the year 2004.
* Though Africa region shows the highest number, the overall highest number of deaths occurred in India which is located in the South-East Asia region.
* All regions in the world show a decrease in number after 2004 except for Eastern Mediterranean, where the number increased at a steady rate of around 750 per year starting from 2000 till 2021.
* The number of deaths due to HIV related causes around the world decreased by around 64 percentage from 2004 to 2021.

# References

* <https://www.who.int/data/gho/data/indicators/indicator-details/GHO/number-of-deaths-due-to-hiv-aids>
* The AIDS Epidemic in 2004 by Robert Steinbrook, M.D.