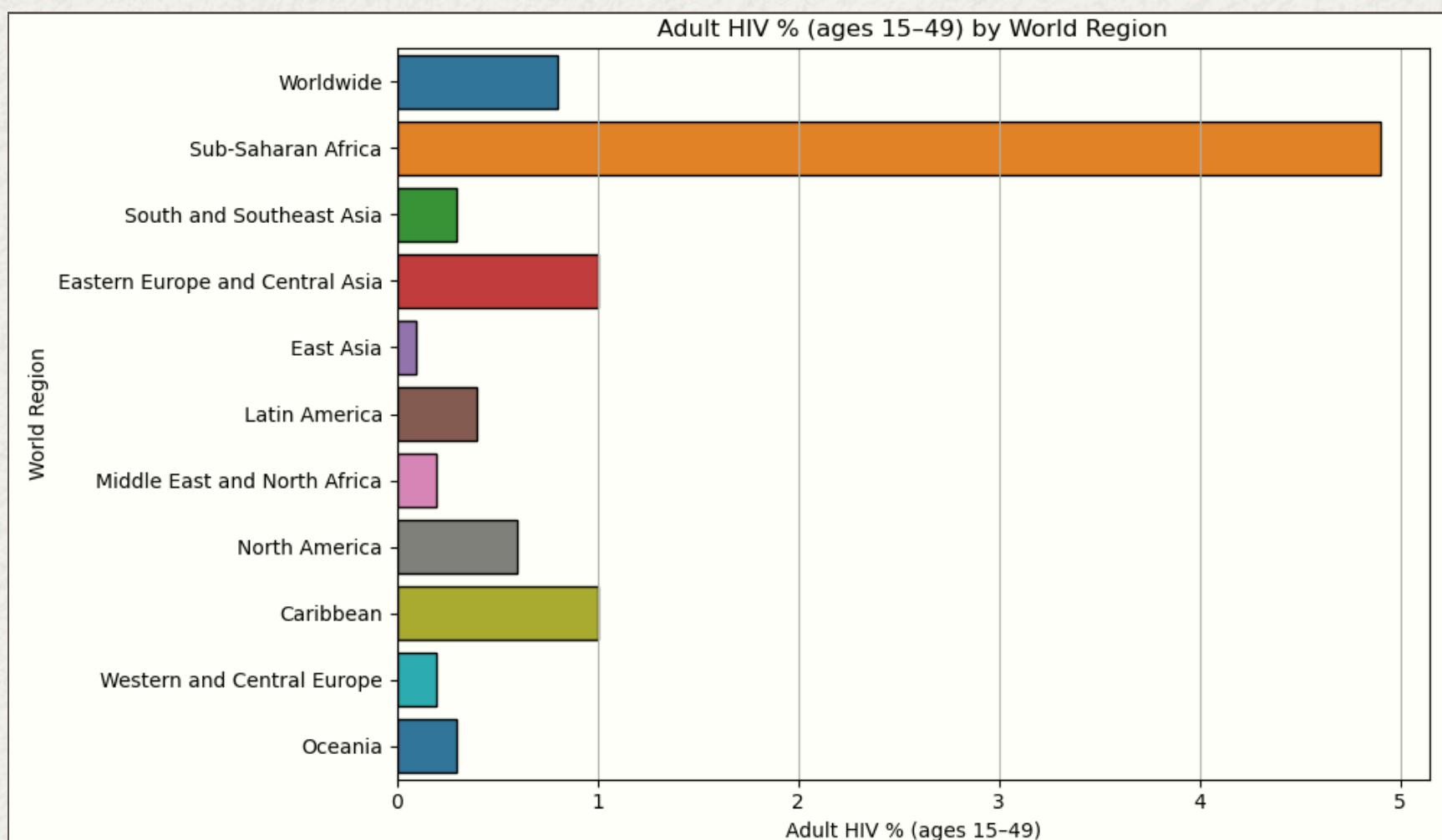


# HIV/AIDS Data Unveiled

In this data-driven blog post, we delve into the prevalence of HIV/AIDS worldwide, shedding light on its significance and the imperative to disseminate information on this crucial topic. By examining where the disease is most prevalent globally, we gain insights into the scope of the issue and potential challenges associated with data interpretation.

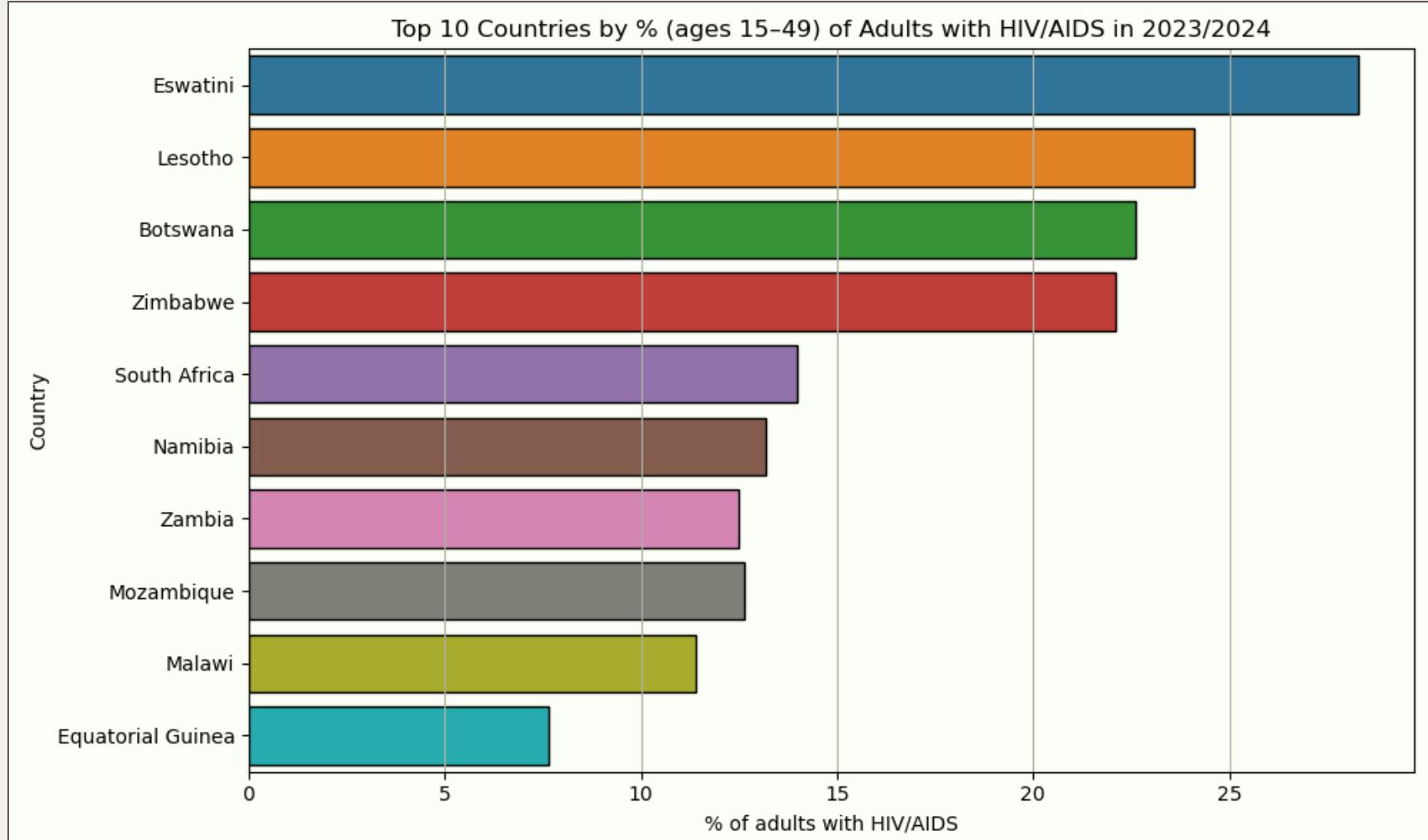
## World Wide Data



Data Web Scrapped from:  
[https://en.wikipedia.org/wiki/HIV/AIDS\\_in\\_Africa](https://en.wikipedia.org/wiki/HIV/AIDS_in_Africa)

HIV/AIDS remains a pressing global health concern, with profound implications for individuals and communities worldwide. The provided graph on above illustrates the percentage of adults aged 15-49 living with HIV across different regions. Notably, Sub-Saharan Africa emerges as the epicentre of the epidemic, demonstrating the highest prevalence rates by a large margin. Additionally, regions such as the Caribbean and Eastern Europe, along with Central Asia, exhibit significant burdens of the disease.

## Sub-Saharan Data



Data Web Scrapped from:

[https://en.wikipedia.org/wiki/List\\_of\\_countries\\_by\\_HIV/AIDS\\_adult\\_prevalence\\_rate](https://en.wikipedia.org/wiki/List_of_countries_by_HIV/AIDS_adult_prevalence_rate)

Exploring further into Sub-Saharan Africa, we uncover alarming statistics that shows the severity of the HIV/AIDS crisis within this region. Countries such as Eswatini, Lesotho, and Botswana bear the highest occurrence of the disease, with prevalence rates ranging from around 22% to nearly 30%.

Investigating the factors contributing to this disproportionate impact is vital, to understand what trends within a country can lead to these infection levels. We will take a deeper look into the distributions and relationships surrounding HIV/AIDS in Southern Africa later on in this blog.

## Collecting the Data

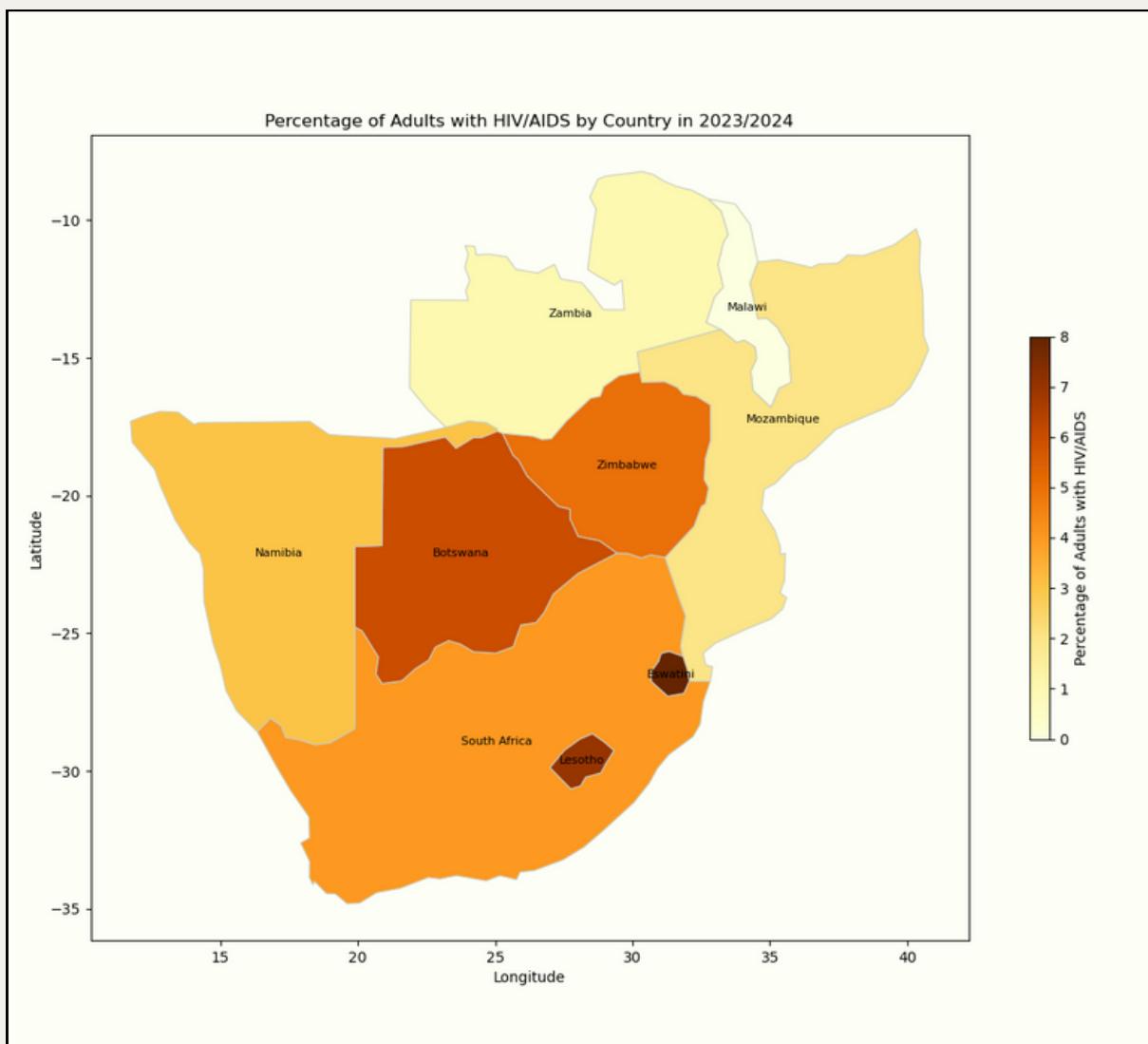
To collect data for the visualisations on this page, I employed web scraping techniques to extract information from the relevant Articles and Wikipedia pages. Using a popular coding language called Python, I created scripts to ‘scrape’ or ‘read’ the web pages and store the data in tables to allow me to analyse the underlining trends. After cleaning the data (checking for missing values and removing unwanted symbols), I was able to create the visualisations you see throughout this whole blog post.

## Economic Factors in Southern Sub-Saharan Africa

To ensure the data was accurate for the tasks at hand I compared the list of countries missing from the dataset to a list of Sub-Saharan Countries and found that there were only 3 missing entries. ‘Somalia’ is the largest on these and most likely to cause miss leading trends however, like the other two, it is not near the countries I intend to look at for the rest of the blog and should not have a significant impact.

Missing Countries
Somalia
Comoros
Seychelles

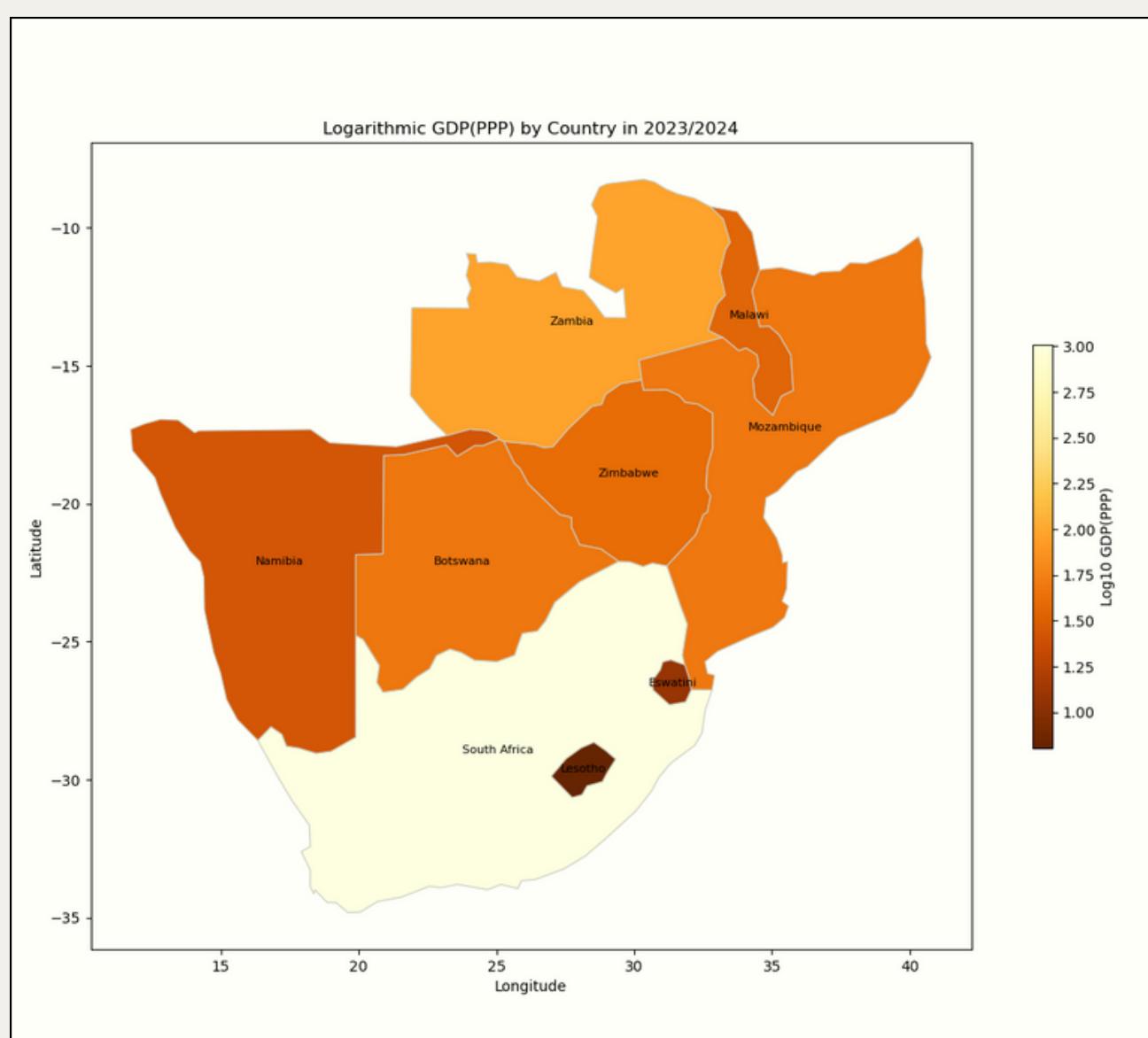
## Mapping HIV/AIDS in Africa



Data Web Scrapped from:  
[https://en.wikipedia.org/wiki/List\\_of\\_countries\\_by\\_HIV/AIDS\\_adult\\_prevalence\\_rate](https://en.wikipedia.org/wiki/List_of_countries_by_HIV/AIDS_adult_prevalence_rate)

In the visualisation above, a physical map of Southern Africa highlights the distribution of HIV/AIDS prevalence among countries, with darker shades indicating higher percentages of adults affected by the disease. A striking observation emerges as we notice the geographical proximity of Eswatini and Lesotho, both situated within South Africa, yet exhibiting stark differences in infection rates. This phenomenon prompts further investigation into the underlying factors contributing to such disparities within neighbouring regions. Additionally, a discernible trend emerges as we move northward into Africa, with countries showing progressively lower prevalence rates. However, it's crucial to acknowledge the potential influence of data availability and research efforts on these observed patterns. Particularly in more remote areas, where research funding may be more limited.

## Mapping GDP(PPP) in Africa

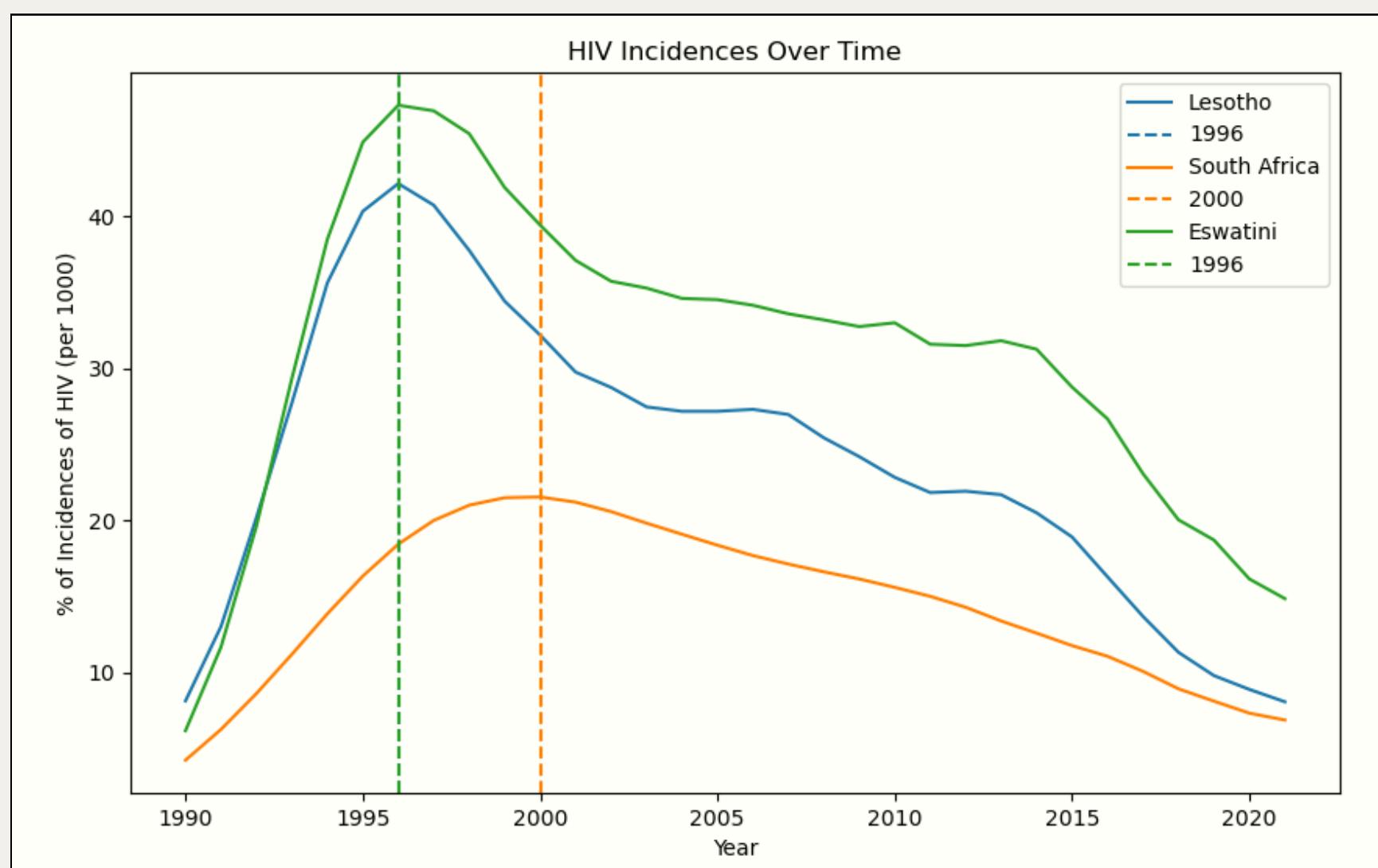


Data Web Scrapped from:  
[https://en.wikipedia.org/wiki/List\\_of\\_African\\_countries\\_by\\_GDP\\_\(PPP\)](https://en.wikipedia.org/wiki/List_of_African_countries_by_GDP_(PPP))

Shifting our focus to the new visualisation above, the same map is presented, but this time, darker shades indicate countries with lower Gross Domestic Product (GDP) based on Purchasing Power Parity (PPP). GDP (PPP) represents the total value of goods and services produced within a country adjusted for differences in price levels compared to other countries, providing a more accurate comparison of economic performance. The analysis of GDP (PPP) in relation to HIV/AIDS infections reveals intriguing trends.

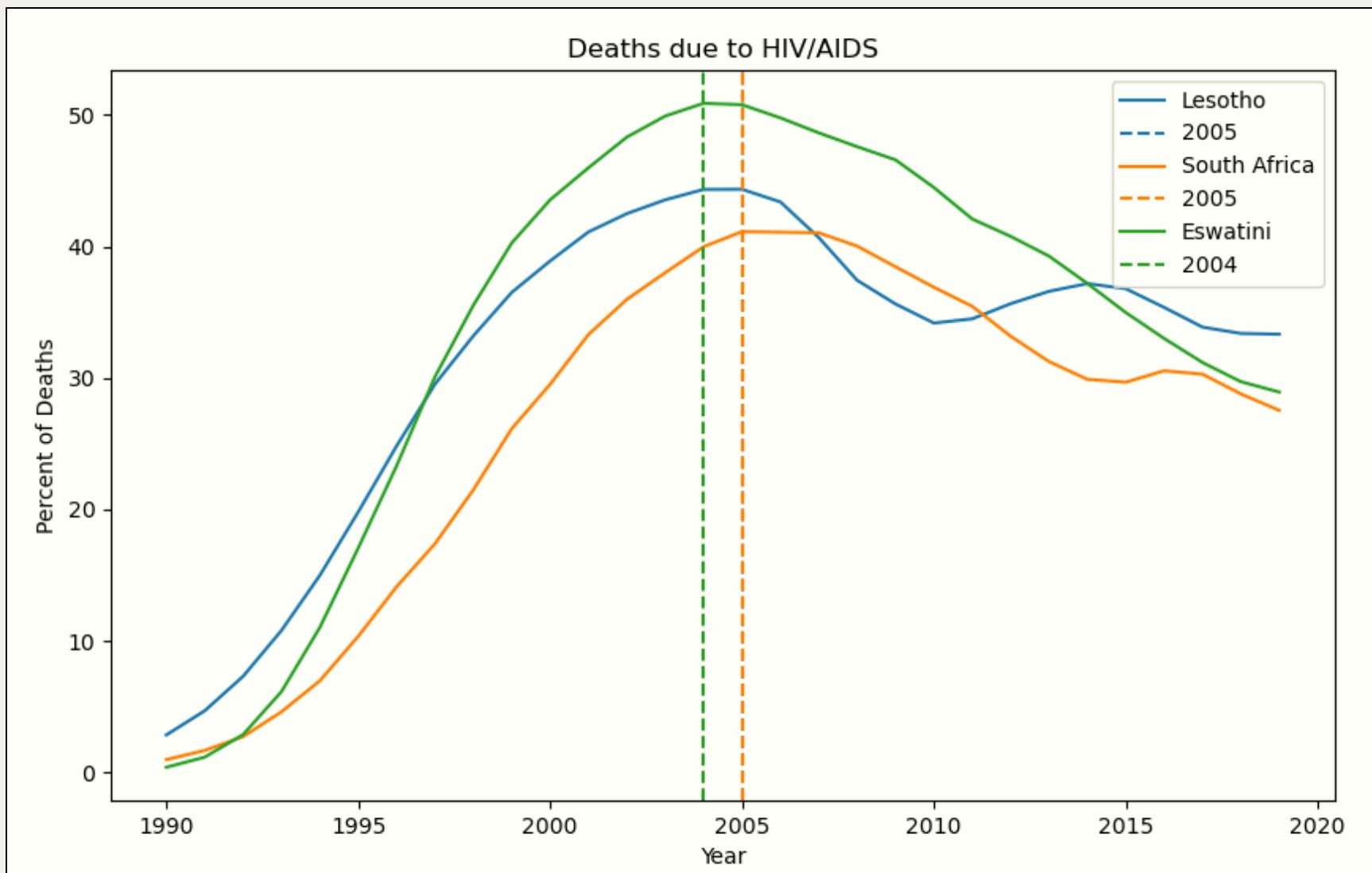
The two maps show that the countries with lower GDP (PPP) exhibit much higher prevalence of HIV/AIDS. This seems to be irrespective of their neighbouring countries and their own infection rates, reflecting the intricate interplay between economic factors and health outcomes. This correlation emphasises the importance of socioeconomic factors in shaping the prevalence and impact of HIV/AIDS within communities and nations.

## A Closer Look at Lesotho, Eswatini, and South Africa



Analysing the map plots, particularly focusing on Lesotho, Eswatini, and South Africa, reveals intriguing patterns worth exploring further. The plot above depicts the trends in HIV prevalence from 1990 to 2020 sourced from the World Data Bank.

Distinctly, we see a peak in HIV incidences in both Eswatini and Lesotho in 1996 and a peak in South Africa in 2000. Following these spikes there is a general downwards trend in cases as healthcare innovations improved. Moreover, we observe the impact that GDP (PPP) has on HIV incidences. This graph consolidates the link between the two, with countries with higher GDP (PPP) consistently showing lower HIV cases.



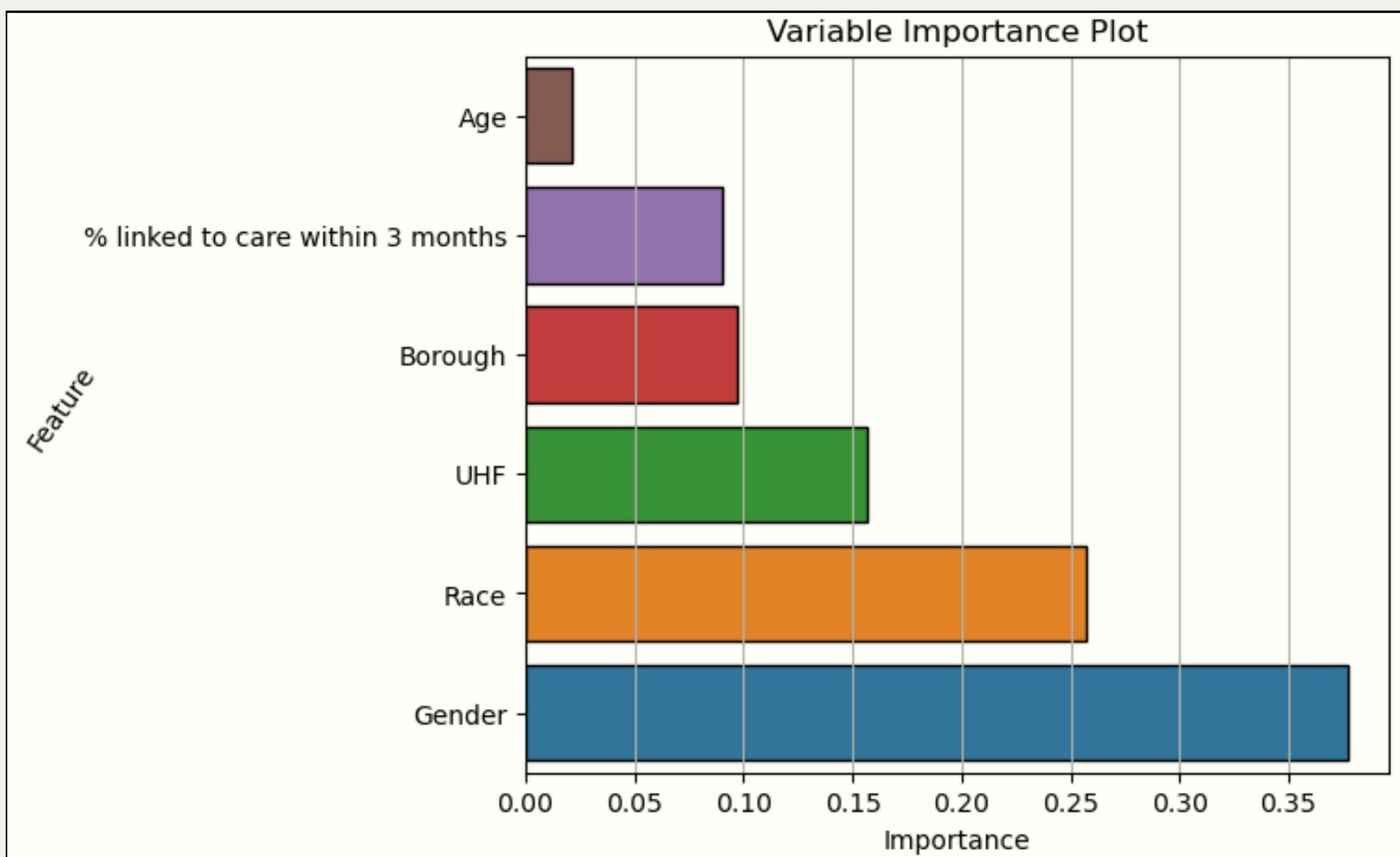
Data Collected from:  
<https://ourworldindata.org/hiv-aids>

A striking observation is the delay in the manifestation of the deadly symptoms of HIV/AIDS, particularly evident in economically poorer countries. This delay is also reflected in the onset of AIDS-related complications, suggesting a lag between HIV infection and the development of severe health consequences. Specifically, our analysis unveils a notable spike in HIV incidences around 1996, followed by a corresponding surge in deaths in 2004, particularly pronounced in Eswatini and Lesotho.

These findings align with the hypothesis put forth by the World Health Organization (WHO), which suggests that individuals with HIV typically exhibit signs of HIV-related illness within 5–10 years of infection. Our data and accompanying graphs provide empirical evidence supporting this hypothesis, as evidenced by the peak in deaths occurring precisely within the anticipated range of 5–10 years following the peak in incidences. This temporal alignment underscores the predictable progression of the HIV/AIDS epidemic and underscores the importance of timely interventions and healthcare strategies to mitigate its impact.

By juxtaposing data from multiple sources and drawing correlations between HIV incidence, mortality rates, and economic indicators, we gain a more comprehensive understanding of the complex dynamics driving the spread and impact of HIV/AIDS in affected regions. These insights are invaluable for informing targeted interventions and policies aimed at addressing the multifaceted challenges posed by the epidemic.

# Contributing Factors to AIDS Diagnoses



Data Collected from:

<https://data.world/login?next=%2Fcity-of-ny%2Fju2dad%2Fworkspace%2Ffile%3Ffilename%3Ddohmh-hiv-aids-annual-report-1.csv>

Displayed here is a variable importance plot revealing the key attributes influencing AIDS diagnoses.

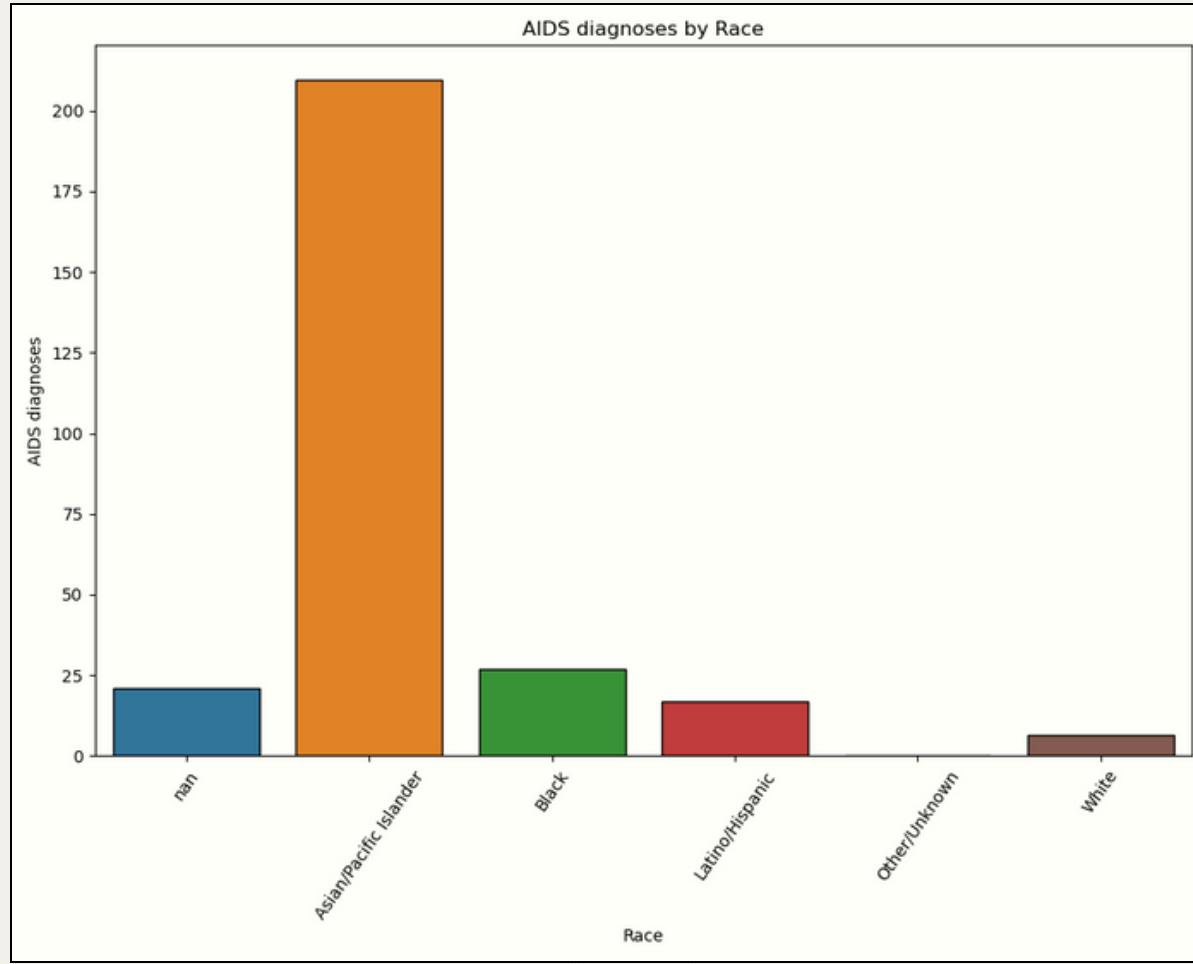
For this section I decided to utilise a Random Forest as I have some prior experience using this method. To sum up quickly, the code takes lots of differing variables and is able to rank them on how much of an effect they have on a desired outcome. For example in this case we are looking at factors that lead to an Aids Diagnosis.

This data originates from New York state, acknowledging the potential presence of inherent biases. However, it remains the most extensively funded and reliable publicly accessible dataset at our disposal. Gender emerges as the predominant factor, exerting the most significant influence on AIDS diagnoses, followed closely by race. These are incredibly useful attribute to be aware of as they allow for governments to target their funding and help directly to the ones who need it most.

Notably, the United Hospital Fund (UHF), intricately linked with borough demographics, exhibits a substantial impact on the likelihood of an AIDS diagnosis. It's intriguing to note that age appears to have the least influence on AIDS diagnoses according to this analysis. Further investigation is imperative to uncover nuanced trends and underlying dynamics within this observation. This insight prompts deeper exploration into the intricate interplay of demographics and health outcomes, shedding light on critical aspects of AIDS epidemiology and potentially informing targeted interventions and policies.

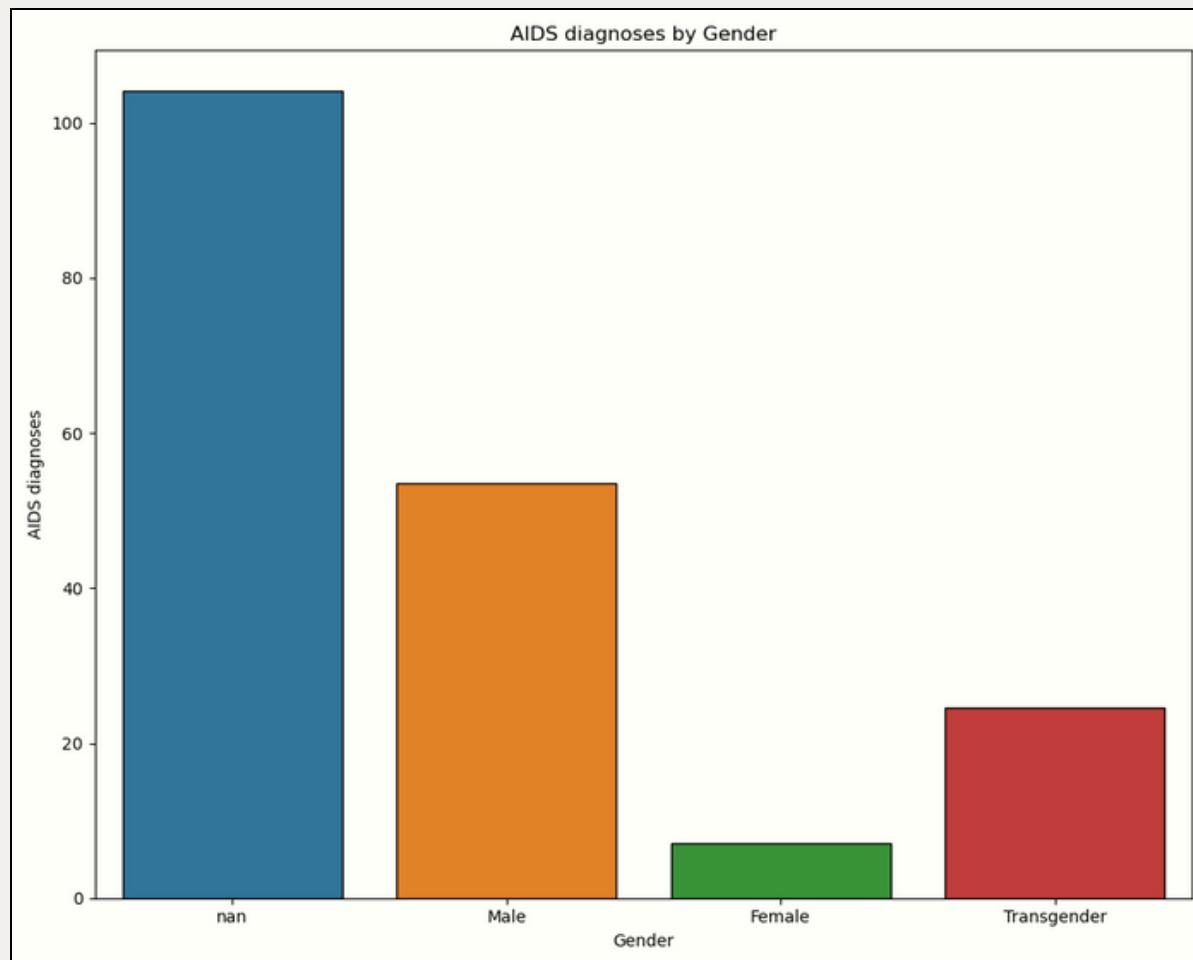
## Deeper Analysis

In this series of three stratified graphs, each delineating distinct demographic segments pertinent to AIDS diagnoses, we embark on an analytical journey elucidating critical insights into the disease's epidemiology.



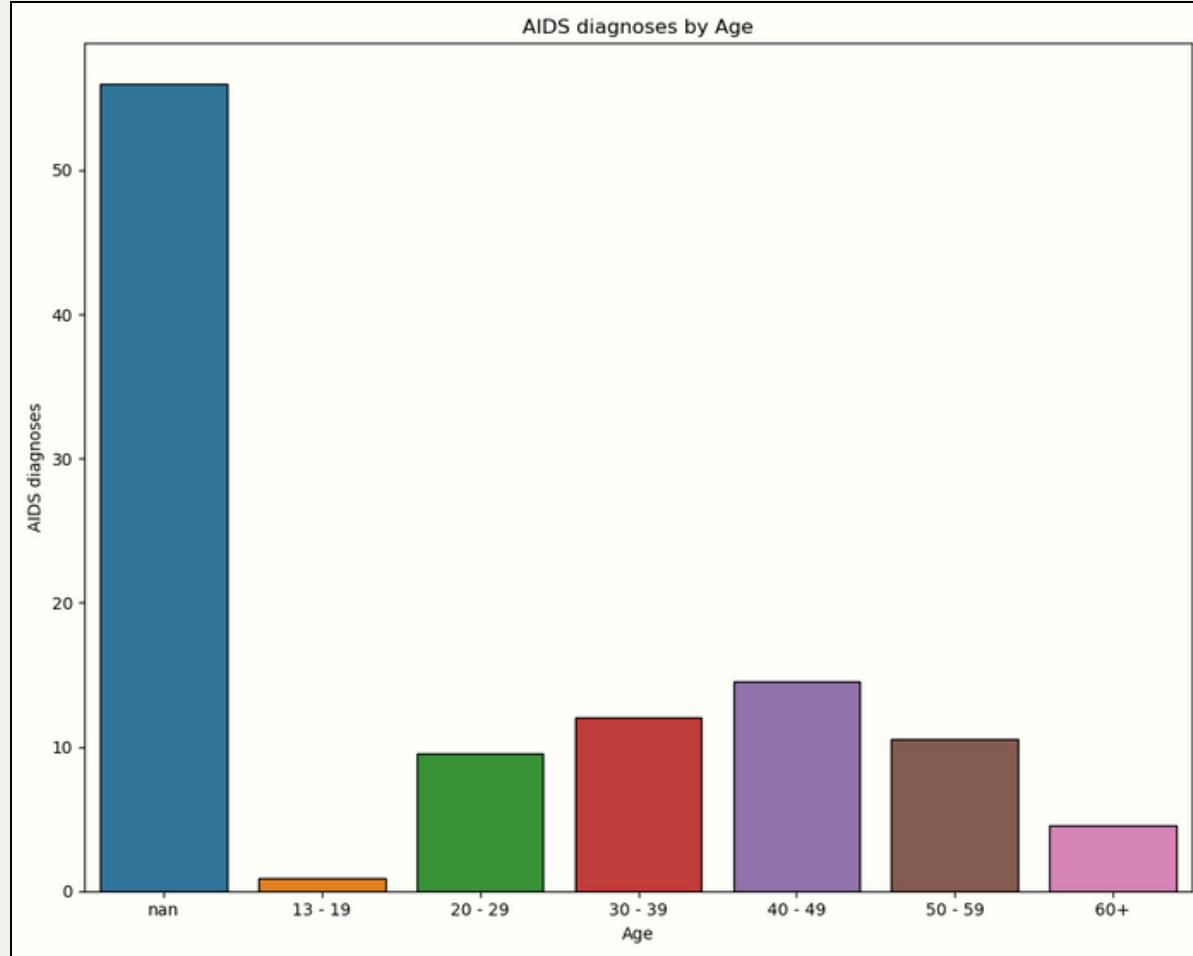
Data Collected from:  
[https://data.world/login?next=%2Fcity-of-ny%2Fju2dad%2Fworkspace%2Ffile%3Ffilename%3Ddohmh\\_hiv\\_aids\\_annual-report-1.csv](https://data.world/login?next=%2Fcity-of-ny%2Fju2dad%2Fworkspace%2Ffile%3Ffilename%3Ddohmh_hiv_aids_annual-report-1.csv)

The first graph unveils a notable surge in AIDS diagnoses within the Asian/Pacific Islander demographic - a noteworthy finding warranting deeper investigation to discern the underlying drivers of this pronounced increase.



Data Collected from:  
[https://data.world/login?next=%2Fcity-of-ny%2Fju2dad%2Fworkspace%2Ffile%3Ffilename%3Ddohmh\\_hiv\\_aids\\_annual-report-1.csv](https://data.world/login?next=%2Fcity-of-ny%2Fju2dad%2Fworkspace%2Ffile%3Ffilename%3Ddohmh_hiv_aids_annual-report-1.csv)

Transitioning to the second graph, a discernible skew towards male individuals in AIDS diagnoses is clear, followed by transgender individuals—a pattern that resonates with established epidemiological expectations. Notably, these trends maintain their significance despite the absence of a substantial portion of age-related data.



Data Collected from:  
<https://data.world/login?next=%2Fcity-of-ny%2Fju2dad%2Fworkspace%2Ffile%3Ffilename%3Ddohmh-hiv-aids-annual-report-1.csv>

Subsequently, our exploration extends to the third graph, delving into the distribution of AIDS diagnoses across various age cohorts. While the variable importance analysis suggests a muted correlation between age and AIDS diagnosis, closer scrutiny reveals a critical caveat: the prevalence of null data obfuscating substantive insights. Upon mitigating this null data, a nuanced binomial bell curve appears, peaking within the 40–49 age range.

While the current dataset may exhibit limitations, these preliminary glimpses offer valuable insights, hinting at broader trends awaiting illumination through more exhaustive and robust data collection methodologies.

## Conclusion

In conclusion, our journey through the data-driven exploration of HIV/AIDS has illuminated critical facets of this global health crisis. From the macroscopic view of global prevalence to the microscopic analysis of demographic influences, each insight has contributed to a deeper understanding of the epidemic's complexities.

It is clear that HIV/AIDS remains a formidable challenge with far-reaching implications for individuals, communities, and societies worldwide. Yet, amidst the sobering statistics and daunting challenges, there is hope.

Hope lies in the power of data to inform targeted interventions and policies, in the resilience of communities to unite against adversity, and in the dedication of healthcare professionals and researchers to push the boundaries of knowledge and innovation.

Our journey doesn't stop here. It's a reminder to keep moving forward, to fight for fair healthcare access, and to support those affected by HIV/AIDS.

Together, let's work to create a future where HIV/AIDS is no longer a global problem, but a challenge we've overcome through unity, compassion, and determination.