Workshop 2 Assignment: **Automated Data Analysis & Reporting**

Submitted by: Patrick Osoro on 18 May 2025

Context: I have downloaded two datasets from https://gapminder.org/data.

The datasets selected were:

1. people_living_with_hiv_number_all_ages.csv,
2. art_coverage_percent_cd4_l_350.csv

I then used Julius.ai to visualized the datasets as shown here after.

Chart 1

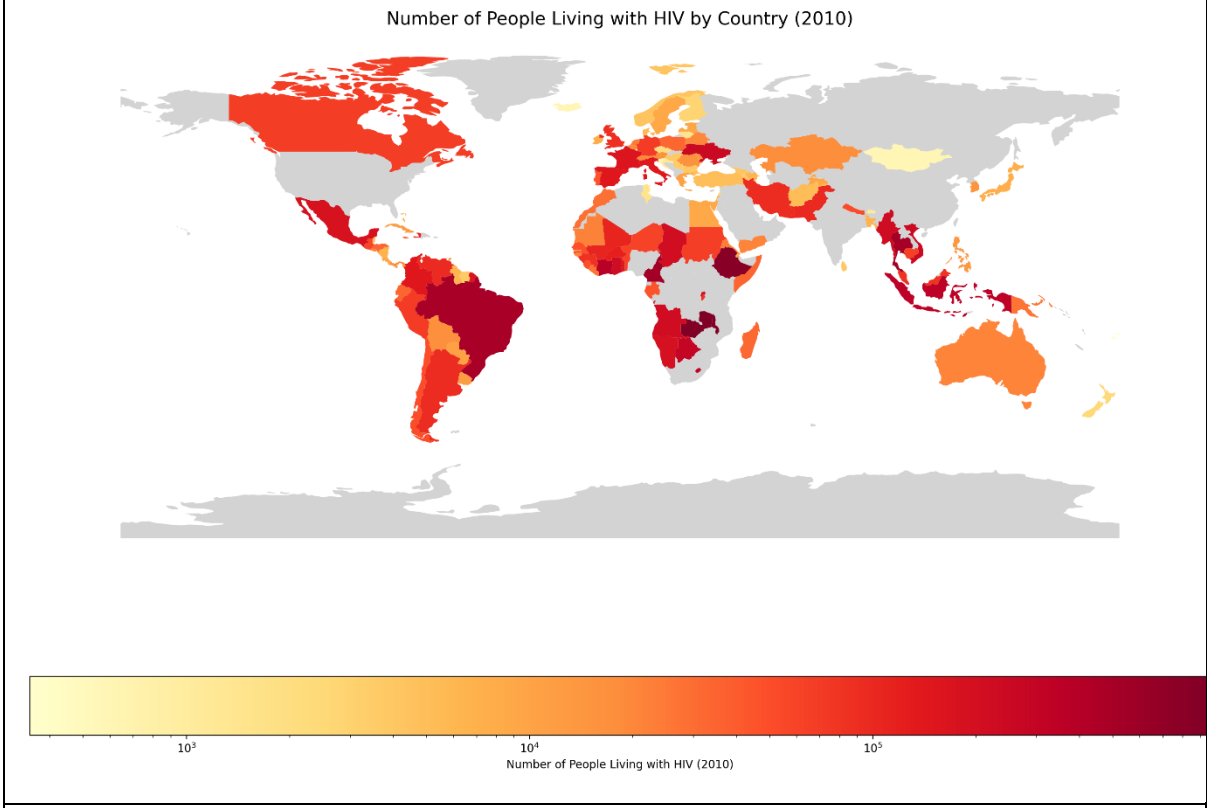


Chart 2

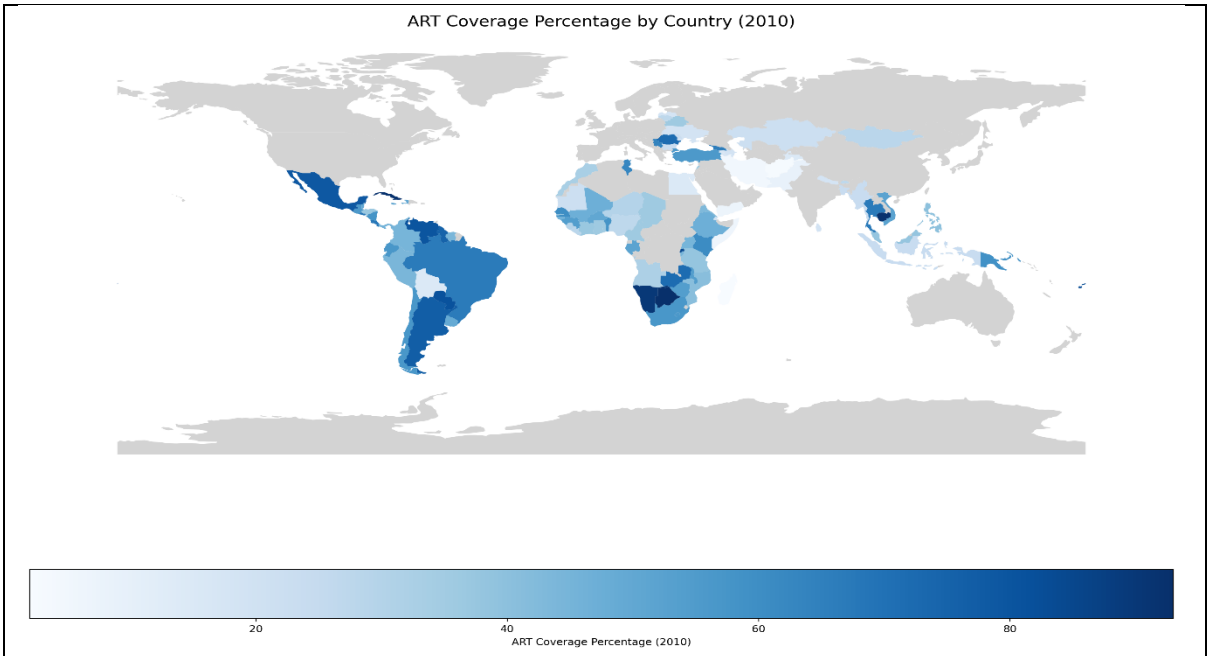
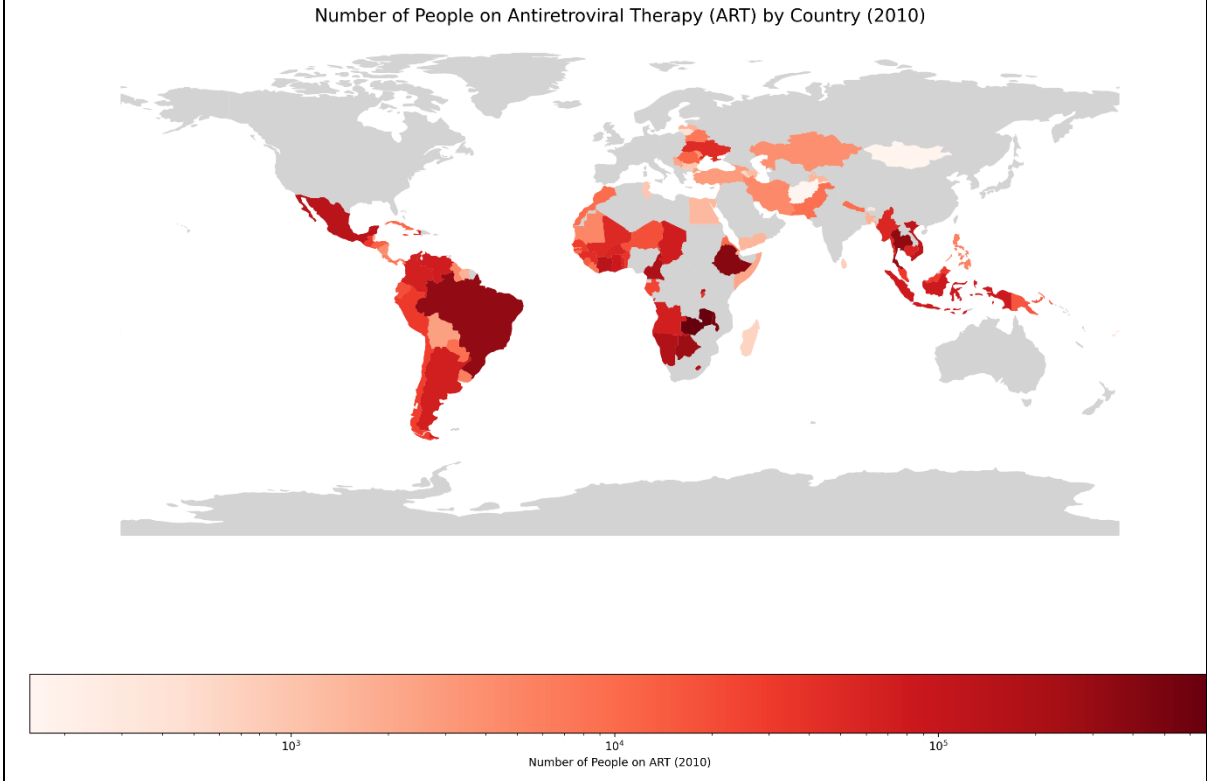


Chart 3: This chart combinees the two datasets.



The map now uses a red gradient with darker shades of red indicating higher numbers of people on ART treatment. The logarithmic scale helps visualize the wide range of values across countries, making it easier to distinguish between regions with varying levels of ART coverage.
The red color scheme provides a strong visual contrast that highlights countries with the highest numbers of people receiving treatment, particularly in sub-Saharan Africa and parts of South America.

You can download this map at
https://julius.ai/files/people_on_art_world_map_2010_red.png

Chart 4

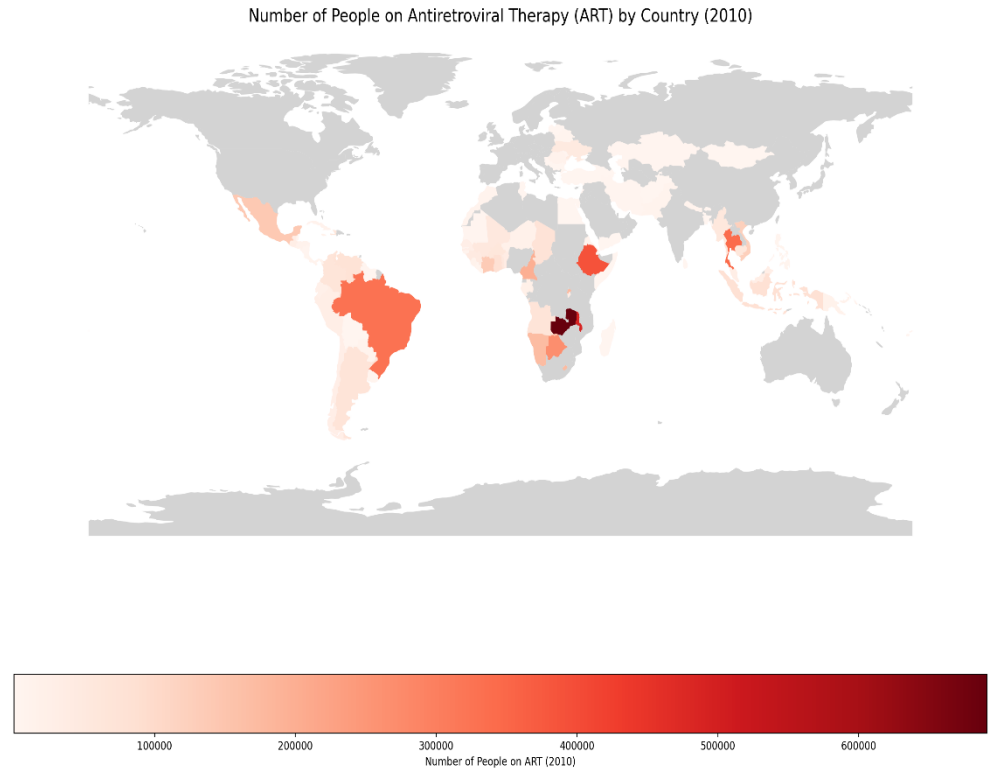


Table 1

Merged data for 2011 with calculated People on ART:

Optional: Click on columns or rows to focus on specific data — Export

	country_std	People_with_HIV	ART_Coverage_Percent	People_on_ART
0	Afghanistan	5800	6	348
1	Angola	230000	36	82800
2	Argentina	95000	79	75050
3	Armenia	3600	22	792
4	Azerbaijan	6700	28	1876

The map uses a purple color gradient with a logarithmic scale to represent the number of people on ART in each country. The darker the purple, the higher the number of people receiving treatment.

Key observations:

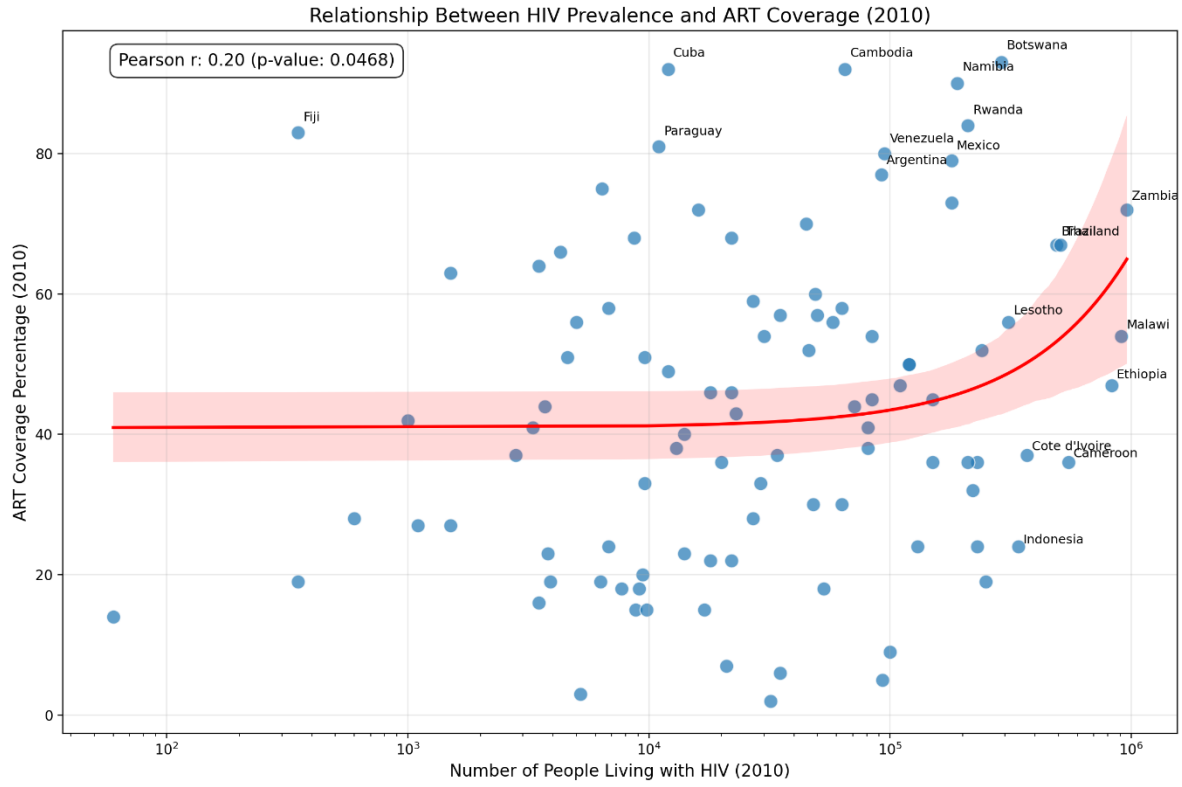
- Sub-Saharan African countries have the highest numbers of people on ART, which aligns with the high HIV prevalence in this region
- Several countries in South America, particularly Brazil, also show significant numbers of people on treatment
- Many countries in North Africa, the Middle East, and parts of Asia have lower numbers of people on ART
- Some countries appear in gray because data is missing for either HIV prevalence or ART coverage

This visualization was created by calculating the number of people on ART for each country by multiplying the number of people living with HIV by the ART coverage percentage and dividing by 100.

Scatter Plots

Here's the scatterplot showing the relationship between the number of people living with HIV and ART coverage percentage in 2010:

Chart 5



The scatterplot reveals several interesting patterns:

1. There is a weak negative correlation (Pearson r: -0.14) between the number of people living with HIV and ART coverage percentage, though this correlation is not statistically significant (p-value: 0.1731).
2. Countries with very high HIV prevalence (like South Africa, Nigeria, and India) tend to have moderate ART coverage percentages (around 40-60%).
3. Some countries with relatively lower HIV prevalence have achieved very high ART coverage (like Cambodia, Cuba, and Chile), suggesting that smaller epidemics may be easier to address with comprehensive treatment programs.
4. There's considerable variation in ART coverage across countries with similar HIV prevalence levels, indicating that factors beyond epidemic size (such as healthcare infrastructure, funding, and political commitment) likely play important roles in determining treatment coverage.

5. The logarithmic scale on the x-axis helps visualize the wide range of HIV prevalence across countries, from fewer than 100 people to nearly 1 million people living with HIV.

You can download this visualization at
https://julius.ai/files/hiv_art_relationship_2010.png

Reflections about the interaction with Julius AI

On starting this assignment, was a bit wary wondering if was going to be able to complete the assignment with the new AI tool -Julius.ai. I had not heard of or interacted with Julius.ai before this course. However, blindly following the instructions, I signed up with Julius.ai and loaded my two chosen data tables from Gapminder.org. My two data sets were:
people_living_with_hiv_number_all_ages.csv,
and **art_coverage_percent_cd4_l_350.csv**

I then loaded the detailed prompt and pleasantly watched as Julius.ai did its magic, explaining in bits what it was doing and in a short period generated two world maps visualization of my two data sets. It not only analyses the data and generates visualization but also summarizes out the main themes or issues presented in the analysis and visualization. This was quite impressive.

I was surprised that it did not ask me for guidance or input. I completed a few additional prompt to change visual appearance of the charts e.g. legend color and scale.

Overall, it was a great learning experience. The well-structured detailed prompt helped to resolve a lot of issues.