*You have three (3) hours to complete the tasks in this test.*

Using the R programming language and Microsoft PowerPoint, please complete the following 3 tasks.

You are permitted to use any packages that you are confident in, but we recommend using “tidyverse” packages for most of the questions. There is no expectation that you will know how to do all the coding in this test without using other resources to look up new packages, methods, functions, etc. – please use whatever resources you would normally use when coding something new (Google, StackOverflow, ChatGPT, etc.)!

Please use a new R script for each question. Your submission should comprise 3 R scripts and 1 PowerPoint file.

You will need to use the Application Programming Interface (API) of the Global Fund and World Bank to get the data to answer these questions.

* The Global Fund API documentation can be found at [data-service.theglobalfund.org](https://data-service.theglobalfund.org).
* The World Bank API can be accessed through the “WDI” R package or the data can be accessed directly on the World Bank website at [data.worldbank.org](https://data.worldbank.org)

**Task 1**

Using the attached data file “HIVDiseaseBurden.csv”, create a function that does the following:

* Takes as input a number “N” representing total funding dollars, which can be modified by the user.
* The function should then distribute the funding to all countries in the data file according to their share of people living with HIV (disease burden).
* Your function should ensure that no country gets less than $500,000, if a country-disease gets less than $500,000 then it should be increased by taking proportionally from all other countries.
* Your function should ensure that no country gets more than 10% of the total funding. If a country is in breach of this limit it should be reduced to 10% of the total funding and the amount in excess should be redistributed to other countries proportionally with their disease burden.
* The output of the function should be a dataframe containing (at least, but not limited to) the following columns: ISO3, DiseaseBurden, AllocatedAmount, where allocated amount is the amount of funding for each country resulting from the distribution described above.
* Tips to check your output:
  + Sum of the AllocatedAmount column should be “N”
  + No value in the AllocatedAmount column should be less than 500,000
  + No value in the AllocatedAmount column should be more than N / 10.

**Task 2**

From 2014-2022, Global Fund allocated ~32% of resources to malaria. **Produce no more than two slides arguing for an increased share of Global Fund funding to be allocated to malaria**. Each slide should contain at least one graph produced in R and some concise explanatory text.

Among the areas you may want to consider are:

* The epidemiological situation of each disease globally
* Impact and progress against the three diseases in recent decades
* The relative economic capacity of states where each of the three diseases are concentrated
* The impact on any current Global Fund investments.

You do not need to address all of these areas nor are you limited to only analyzing these areas; your focus above all should be on constructing a persuasive argument using data.

**Task 3**

Using the Global Fund’s API as a data source, make one treemap plot that shows the recent Global Fund investments in each country and disease (note that a country may have multiple grants in the same disease program). The area of each box in the treemap should correspond to the size of the investment.