SEMESTER/YEAR : **TRIMESTER IV/2024**
SESSION : **Mar24 – Jun24**
COURSE CODE : **TBD**
TITLE OF THE COURSE : **INTRODUCTION TO ANALYTICS**
FACULTY MEMBER : **ARVIND VENKATADRI**

Dayananda Sagar University Bengaluru

ASSIGNMENT #1

Instructions

Task

1. Look at the graph in the picture:

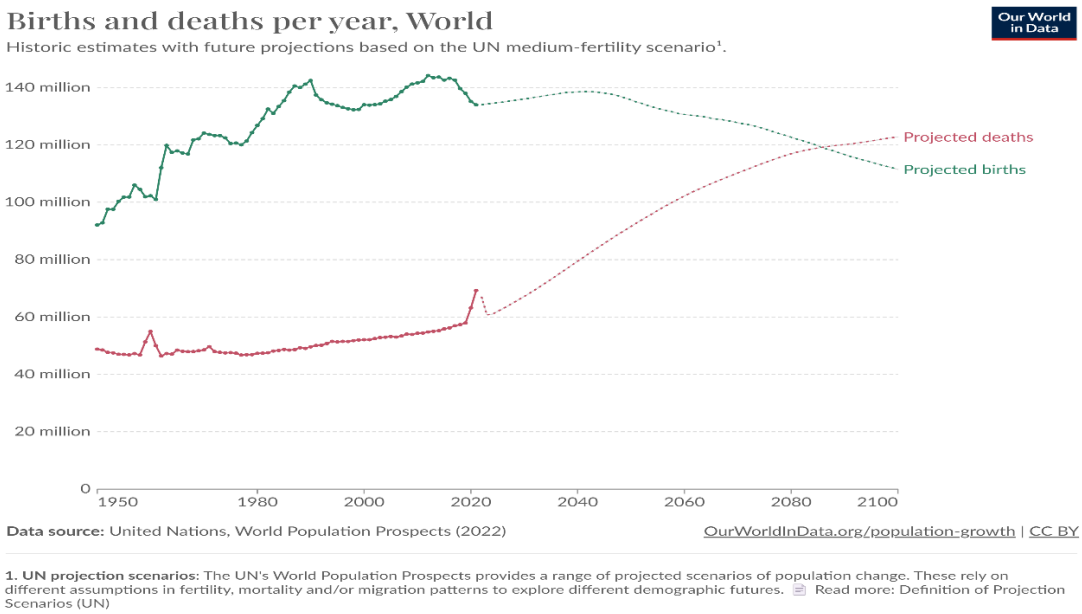

Figure 1: https://ourworldindata.org/grapher/births-and-deaths-projected-to-2100?time=earliest..2100

1. Your task is to ***reproduce this picture in R and explain it*** in a well-crafted HTML report created using RStudio/Quarto.
2. The CSV data is available here: https://ourworldindata.org/grapher/births-and-deaths-projected-to-2100?tab=table&time=earliest..2100

Workflow

1. Fire up Rstudio.
    a. Click on the "Project" button in the upper-right corner.
    b. Click on New Project -> New Directory -> Quarto Project.
    c. Ensure you name the Project "Your_Name_A1". E.g "Arvind_A1".
    d. RStudio will create a new project within your default "R_Work" folder AND give you a ready-made Your_Name_A1.qmd file. This is the file you will edit to create your graph.
2. Set up the YAML as needed for HTML report creation.
3. Load the R packages you need: mosaic, tidyverse, ggformula and others from class. Do not use R packages not introduced in class. Do not use Base R commands either; use only tidyverse commands.

4. Create a /data folder inside your project. Download the CSV and save it there. Now, read in the data and rename variables as needed.
5. Introduce the data set: its dimensions, variables, what they mean.
6. Recreate the graph in the picture using code in R. Mention and justify any decisions you might have made in creating this graph. I repeat: Do not use R packages not introduced in class.
7. Add explanations and interpretations from the graph. Explain interesting trends, points of interest on the graph.
8. Include code that you tried and which "did not work," with possible reasons; (Use "eval=FALSE" in your code chunk to show up errored code)
9. Add comments as to what you were **not able to achieve** in your code. E.g. the kind of plot, the sizes, or the colours, scales etc. Comment on why this might have happened. And on how it might affect your interpretations.
10. NOTE: your .qmd file MUST render to HTML as a readable report. (That is the first thing I will look for). Check for this BEFORE you submit.

What to Submit

1. A **.zip** file that contains your ***ENTIRE A1 R Project FOLDER***. It should contain your .qmd file, your /data folder with the CSV file, both named appropriately so it will be readable in your code. (There may be other files too, created by R). I will demonstrate this in class.
2. NOTE: I will **not** edit your code. I will simply open your project and run your code and render it. This IS a test of your ability to code, to look through the help files for each R package, and use commands to suit your purpose, and to produce a human-readable report.

Assessment

Total Marks: 30

1. Presentation: Structure of your quarto doc; code chunks; labels, choice of variable names, the Graph; title for graph; colours; annotations; .qmd Rendering; Organization of your Document; File Naming; Data CSV
2. Content: Flow of information in your Quarto doc; Introduction of dataset; Explanation of Graph; Comments in Code; Explanation and insights, Conclusions.
3. Originality:
   - Including code which "did not work," with possible reasons.
   - Is the code your own? Or "leveraged" from the web? (I have ways of finding out ;-D)
   - Use of AI tools like ChatGPT is permitted, but I will ask for an explanation of your code over 1-on-1 review call if I feel that is necessary.
4. Adherence to Deadline:

Delay	Total (out of)
None	30
<=1 wk	26
<=3wk	20
>= 3wk	0

=X=