Technical exercise for Extended Term Consultant position

Context:

The purpose of this exercise is to assess the following:
- Your ability to pull data from an API
- Your ability to conduct some basic data wrangling tasks
- Your ability to convey information through data visualization
- Your ability to think statistically
- Your ability to communicate findings to both technical and non-technical audiences
- Your level of comfort with coding and reproducible research

There are two parts in this exercise. In Part I, you will complete a practical, hands-on exercise. In Part II you will be asked to review and critique a short statistical analysis.

Deliverables:

In total, you will produce <u>three documents</u> and publish them via a Github repository (or another equivalent platform: Gitlab, Bitbucket, etc.)
Your repository should look something like this (Github example):
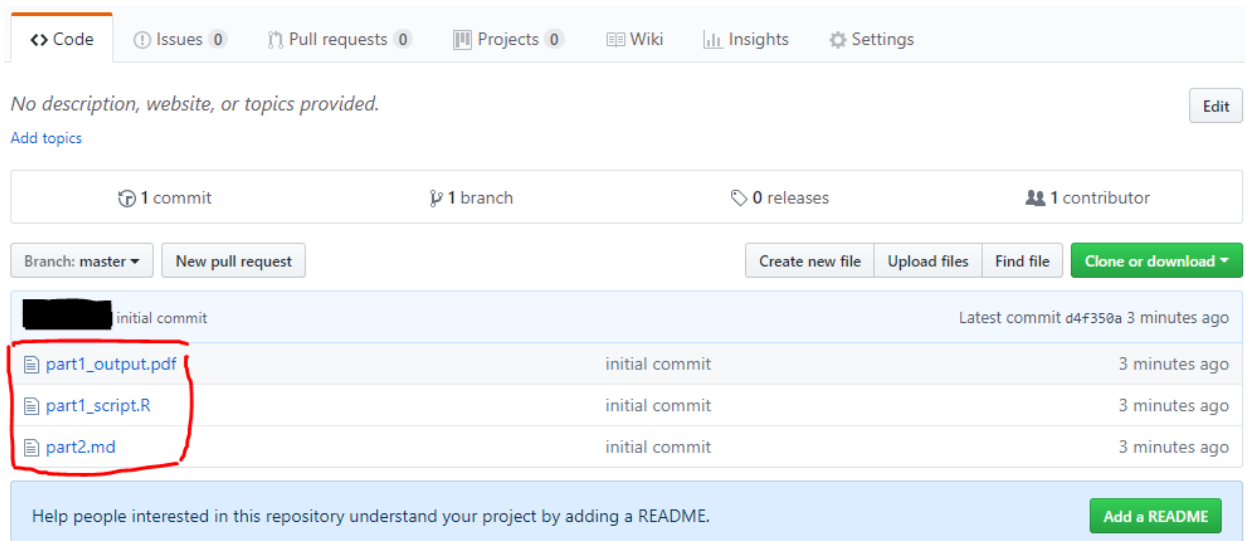
- **Part I - Two documents**
  Your objective is to produce a one pager analysis of improved sanitation facilities (% of population with access). Your one pager should contain **one chart** highlighting your main finding(s), as well as a **written paragraph** that briefly summarizes your approach and finding(s).
  Your work will be reviewed by two different panels: A technical one, and a non-technical one.

You should produce two documents:
- The one pager containing your chart and paragraph. This will be reviewed by the non-technical panel. This document should be in a format that is easy to share via email with a non-technical audience: .docx, .pdf, standalone html file, link to webpage, etc.
- The script that was used to generate the one pager. The script will be reviewed by the technical panel.

   Your code should be well documented and fully reproducible. Your code should run without error on the reviewers' computer, and the output of your code should be the one pager summarizing your analysis (you can use Rmarkdown or Jupyter notebook).

   You can safely assume that the reviewers will have all the necessary software, packages or libraries installed on their computers.

- **Part II - One document**
  A document containing a couple of paragraphs (Word document or similar).

Instructions:

- **Part I**

  We are interested in the following questions:
  - How is access to sanitation facilities changing over time?
  - Is there an overall trend?
  - Does it vary by income group?

    1) Use the World Bank API to retrieve the basic sanitation facilities (% of population with access) (code: SH.STA.BASS.ZS) data for all countries in the WDI starting 1960. Please feel free to use any existing API client (R, Python, Stata)
    2) You most likely won't be able to use the data as is. Conduct any data wrangling tasks you feel are necessary to transform the raw data into a usable format.
    3) Identify one good visual display of the data that will allow a reader to easily answer the three questions mentioned above.
    4) Produce **one** high quality chart that will convey this information.
   5) Summarise your findings in one short paragraph.

The idea here is just to assess your ability to pull data from an API, conduct some basic data wrangling, and make a sensible visual display of the data, so **please don't spend too much time** trying to over-analyze the data.
- **Part II**

You receive the following email from a World Bank colleague. What do you think of the claims made in that email? Please review and critique. Your response should be concise (~300 to 500 words)

How long do World Bank retirees live?
We collected information about all World Bank retirees who died within the past 6 years. We compiled that information, and came up with the following distribution of 919 deceased retirees.
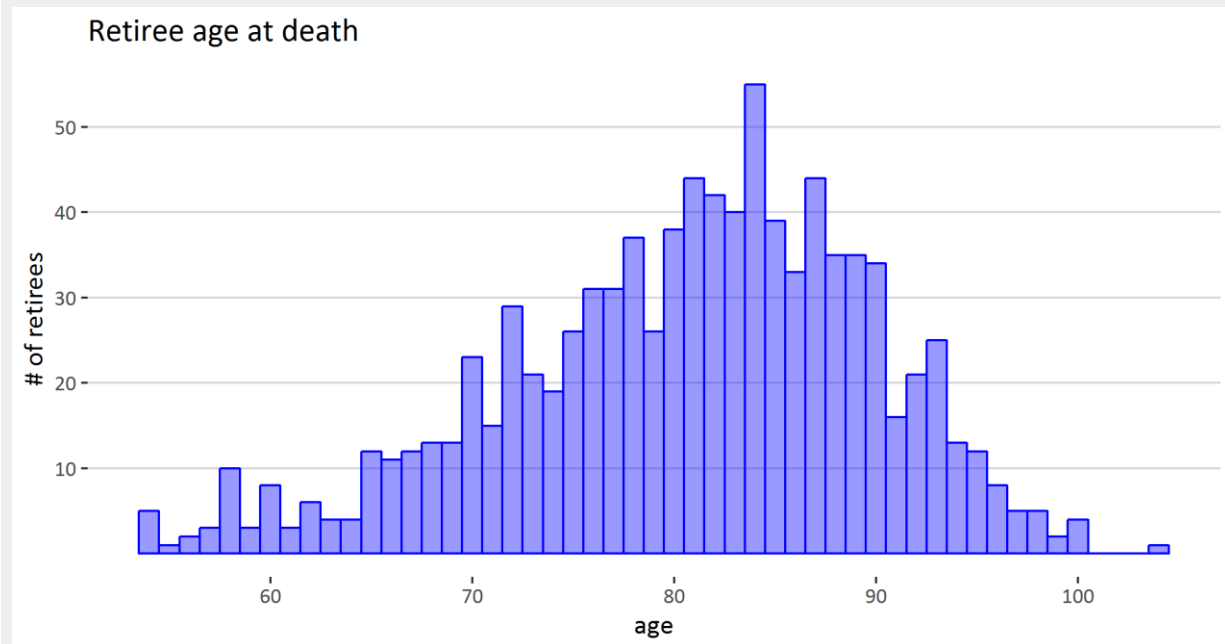
It's a wide distribution. Some were unfortunate enough to expire in their 50s. At the other extreme, one hardy soul lasted until 104!
The mean for this distribution is 81 years old. This is not very impressive: compare it to the US, where a man and woman reaching age 65 today can expect to live, on average, to ages 84.4 and 86.6, respectively.

Please email the URL to your repository to rohagan@worldbank.org by midnight April 23, 2023 (Washington DC time).

Questions and clarifications
For questions and clarifications, email rohagan@worldbank.org by April 21, 2023.