Exercise: **Data Management**

Day 2, Part A

1. Load the Nigeria health metrics data set ('data/nigeria_healthmap.csv').

   a. Keep just the rows where the `geography` variable is 'National'.

   b. Keep just the `year`, `indicator`, `units`, `estimate`, `ci_lb`, and `ci_ub` variables.

   c. Load the population counts for Nigeria and merge onto the health metrics data set ('data/nigeria_pop.csv').

   d. Sort the data by year and then indicator.

   e. Rename the variables `estimate`, `ci_lb`, and `ci_ub` to `est`, `lwr`, and `upr`, respectively.

   f. Save your formatted data as a .rds file in the 'output' folder of your main directory.

2. Load in the Ebola geospatial data (two files: 'data/ebola_point_data.csv', and 'data/ebola_polygon_data.csv').

   a. Combine the point and polygon data into one data frame.

   b. Keep only the `UNIQ_ID`, `Country`, `Virus`, `LAT`, `LONG`, `STR_YEAR`, `OB_CASE`, `OB_DEATH`, and `CASE_TYPE` variables.

   c. Rename these columns to `ID`, `Country`, `Virus`, `Latitude`, `Longitude`, `Year`, `Cases`, `Deaths`, and `Type`, respectively.

   d. Keep only rows that refer to 'index' type cases.

   e. Sort the data by country and year.

   f. Save your formatted data as a .csv file in the 'output' folder of your main directory.