NCD Countdown ReadMe

* Scripts
  + This folder contains all the code required for the NCD Countdown paper
  + ghecausematch.R
    - Takes data from WHO and condenses to only included analyzed NCD4 causes as well as all cause mortality estimates for the years 2000-2019
      * Requires dths\_20age.dta which is not stored in Github as it is far too large(~4,000 MB) for the 100 MB file size limit
    - This script also updates the column names and age/sex IDs to match GBD data
    - This script also calculates the mortality rate per 100,000 from the number of deaths and population in the original format
    - Outputs = ghedata.csv which is stored in the new\_data folder
* Input\_Data
  + This folder has all the original data from the first paper draft
* new\_inputs
  + This folder contains the data required to run the new analysis
  + ghedata2010.csv to ghedata2019.csv
    - Estimates of mortality rates per 100,000 to be used instead of GBD data
    - One age group different from GBD is “1-11 months”
    - No “All Ages” category – but could add this to the ghecausematch.R calcs
    - Has an extra column called “ghecause” that has the numerical ID for the GHE cause
  + GHEtoGBD.xlsx
    - Matches the names of the GBD causes with the corresponding GHE causes
    - Also includes columns for affected and treated fractions for cases where the GHE cause is at a higher aggregate level than the GBD data
      * \*\*When matching GBD causes with GHE causes using the GHEtoGBD.xlsx spreadsheet you therefore will need to multiple the “Mortality reduction” column in the efficacy file by the “affected fraction” column in the GHEtoGBD.xlsx file.
  + Efficacy0625.xlsx
    - Updated intervention efficacy file
    - See above re: note on updating the mortality reduction estimates if switching to GHE causes
    - Two new intersectoral policies were added (5.5 and 5.6) for “Salt policy” and “Trans fat policy” data for these are in the salt\_policy\_effects.csv file and the tfa\_policy\_effects.csv file.
      * + No data for the following countries: ASM, Virgin Islands, Taiwan, Puerto Rico, Palestine, Marina Islands, Guam, Greenland, Bermuda
    - Some interventions were removed as well as some disease targets upon review – but the structure of the input is the same
      * \*\* the issue with the 2.10 ID is no longer relevant as that intervention has been removed.
  + Coverage0621.csv
    - Only the data were updated – no structural changes or implications for code
    - Old interventions were left in this file so there will be more columns than necessary
  + PINandCosts0625.csv
    - No changes to this file, but note\*\* that impairments for heart failure will now have to be pulled from separate GBD input files (heartfailure.csv). Probability will need to write some code to go from the format in the PINandCosts file which is “Heart failure due to myocarditis” to matching this with “Myocarditis” in the heartfailure.csv file. Also heartfailure.csv file only has data for 2019, so \*\*let me know if we need more/different years.
  + Tobaccoandalcoholefficacy6.csv
    - \*\*Filter by NCD4==yes for NCD4 causes