Instructions replication "Half a million excess deaths in the Iraq war: Terms and conditions may apply"

This repository includes a number of folders containing data and code needed to replicate the results as reported in "Half a million excess deaths in the Iraq war: Terms and conditions may apply". To replicate the results, please read the description below on the contents of the individual folders. Some of the scripts will produce output saved in a folder named `output`.

raw_data

Contains a number of original data files. Included are the original replication files of the Hagopian et al. (2013) study,

- `hh_deaths.csv` contains information on the recorded fatalities
- `hh_roster` is the household survey data
- `pop.csv` is the population data for Iraq

Additionally the folder includes a number of files starting with `IRQ_adm1`. These are all part of the shape file needed to produce the choropleth plots (`figures_annex.R`).

data

- `processed.Rdata` which is the output of `re_analysis.R`.

code

This folder includes all the R-scripts needed for the analysis presented in the paper and supplementary material. It includes

- `functions.R`, this script contains a number of functions such as
  1. a bootstrap function,
  2. robust clustered errors, and
  3. functions for standardising the input variables for the regression
- `replication.R`, this script replicates the results from the Hagopian et al. (2013) study and is based on the Python script provided in the original replication files.
- `identify_governorates.R` is a script that identifies the governorates as described in section A.3.1

- `re_analysis.R`, provides the re-analysis of the data as reported in section 2.3 of the paper. Note that this script also processes the data needed for the figures and the regression, which are saved as `processed.Rdata`.
- `bootstraps.R` produces all the bootstrap estimates and should be run after `re_analysis.R`.[1]
- `prep_regression.R` uses `processed.Rdata` as input along with the shape file to prepare the data for the regression analysis.
- `regression.R` fits the regression models as reported in table 2 and A3.
- `figure_paper.R` produces the figures on the main text and should be run after the analysis is done.
- `figures_annex.R` is the code that creates all the figures included in the supplementary material, including the maps.

[1] NB - running this script can take a while; each bootstrap takes about 15 minutes.