

A word of caution! Not all data sets out there are created equal. Remember you are looking for something that meets the criteria outlined in the assignment. Case names (something that is different for every entry) will not work as a variable. You should look for excel files (xlsx or xls) or csv files (which can be opened by excel and saved as xlsx).

1. Health

- (a) CDC Cancers

2. Sports

- (a) Soccer
- (b) Basketball
- (c) Football
- (d) Hockey
- (e) Baseball

3. Public Policy

- (a) Federal Database
- (b) ProPublica (some of these are not free but there are some)
- (c) National Bureau of Economic Research
- (d) A collection of Global Warming Data
- (e) UN Database
- (f) OKC Data Page

4. General Interest

- (a) Lock5 Textbook Data
- (b) Kaggle a Data Science Competition website
- (c) ACLED Curated data sets on conflict around the world
- (d) NASA
- (e) A Collection of Some of my Favorite Cultural Data sets
- (f) ProPublica

5. Data Set Search Tools and a Meta-List

- (a) Google Data Set Search
- (b) 50 Datasets
- (c) R statistical software Data sets
- (d) Teach Data Science