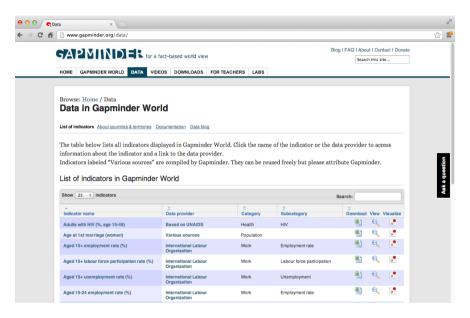# Data resources

Jeffrey Leek
Johns Hopkins Bloomberg School of Public Health
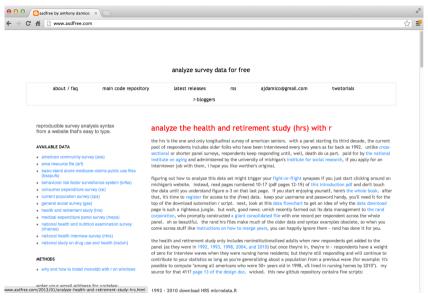
Open Government Sites

- United Nations http://data.un.org/

- U.S. http://www.data.gov/

    - List of cities/states with open data

- United Kingdom http://data.gov.uk/

- France http://www.data.gouv.fr/

- Ghana http://data.gov.gh/

- Australia http://data.gov.au/

- Germany https://www.govdata.de/

- Hong Kong http://www.gov.hk/en/theme/psi/datasets/

- Japan http://www.data.go.jp/

- Many more http://www.data.gov/opendatasites
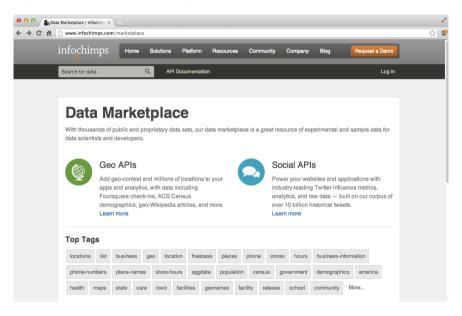
Gapminder

http://www.gapminder.org/

Survey data from the United States

http://www.asdfree.com/

Infochimps Marketplace

http://www.infochimps.com/marketplace

Kaggle

http://www.kaggle.com/

Collections by data scientists

- Hilary Mason http://bitly.com/bundles/hmason/1

- Peter Skomoroch https://delicious.com/pskomoroch/dataset

- Jeff Hammerbacher http://www.quora.com/Jeff-Hammerbacher/Introduction-to-Data-Science-Data-Sets

- Gregory Piatetsky-Shapiro http://www.kdnuggets.com/gps.html

- http://blog.mortardata.com/post/67652898761/6-dataset-lists-curated-by-data-scientists

More specialized collections

- Stanford Large Newtork Data

- UCI Machine Learning

- KDD Nugets Datasets

- CMU Statlib

- Gene expression omnibus

- ArXiv Data

- Public Data Sets on Amazon Web Services

Some API's with R interfaces

- twitter and twitteR package

- figshare and rfigshare

- PLoS and rplos

- rOpenSci

- Facebook and RFacebook

- Google maps and RGoogleMaps