# **Data Sources**

● [Google Dataset Search](https://toolbox.google.com/datasetsearch/)

● [Collection of publicly available datasets](https://github.com/caesar0301/awesome-public-datasets)

● [US Government Open Data](http://www.data.gov/)

● [NYC](https://nycopendata.socrata.com)/[Hoboken](http://www.hobokennj.org/open/)/[Jersey City](http://data.jerseycitynj.gov/) Open Data

● [European Open Data](http://open-data.europa.eu/en/)

● [Yahoo Webscope Program](https://webscope.sandbox.yahoo.com/)

● [Yelp Open Dataset](https://www.yelp.com/dataset)

● [Open data challenges](http://opendatachallenges.org/)

● [GitHub Archive](http://www.githubarchive.org)

● Various Network Datasets:

○ <http://snap.stanford.edu/data/>

○ <https://networkdata.ics.uci.edu/index.html>

○ <http://webscope.sandbox.yahoo.com/catalog.php?datatype=g>

○ <http://www-personal.umich.edu/~mejn/netdata/>

○ <http://socialcomputing.asu.edu/pages/datasets>

○ <https://github.com/gephi/gephi/wiki/Datasets>

○ <http://toreopsahl.com/datasets/>

○ <http://www.casos.cs.cmu.edu/computational_tools/data2.php>

○ <http://deim.urv.cat/~alexandre.arenas/data/welcome.htm>

○ <http://www3.nd.edu/~networks/resources.htm>

○ <http://vlado.fmf.uni-lj.si/pub/networks/pajek/data/gphs.htm>

○ <http://www.crowdflower.com/data-for-everyone>

● [Network Dynamics Datasets](https://dynamics.cs.washington.edu/data.html)

● [Active and completed Kaggle competition](https://www.kaggle.com/competitions)

● [International AAAI Conference on Web and Social Media (ICWSM) 2018 Datasets](https://dataverse.mpi-sws.org/dataverse/icwsm18)

● Active and completed Kaggle competition<https://www.kaggle.com/competitions>. Some of the datasets are not networked by nature, but can be analyzed if you construct a meaningful network.

● Search Google for \_\_\_ data dump. For example, Reddit data dump, StackOverflow data dump