

Data Sources

By Miguel Paz

A curated list of useful databases, open data portals and other selected data sets and tip sheets from various authors.

[Finding Data for Reporting \(selection of resources by Barbara Gray, Associate Professor and Chief Librarian at CUNY J-School\)](#)

[US Census](#)

[Census Fact Finder](#)

[Census Quick Facts \(Good summary of census data points\)](#)

[Census Easy Stats \(fast way to get updated stats about a place\)](#)

[Community Facts \(fast answers about popular data questions\)](#)

[Census Stats \(databases on Building Permits, Counties, Occupation by Race, Hispanic Origin and Sex, International Trade Data\)](#)

[American Fact Finder Advanced Search](#)

[Census Databases and Resources \(CUNY J School\)](#)

[Data.Gov USA \(US Government Data Site\)](#)

[Data.gov JSON sources Government metadata.](#)

[Data USA \(Data profiles from Census on nation, state and city level\)](#)

[Census Reporter \(Data profiles from Census on nation, state and city level with filtering\)](#)

[CUNY Journalism Research Center Research Guides](#)

[NY City The City Record Online \(CROL\)](#): Updated each business day, includes procurement notices (solicitations and awards) from 2003 to present while non-Procurement notices (public hearing and meetings, court notices, auctions and sales, agency rules and special materials) are searchable from August 2013 to present.

[NY State Open Data](#)

[NY City Open Data](#)

[NY City Data 2 Go](#)

[NYC Business Atlas](#) Explore business conditions in NYC neighborhoods

[NYC Citygram](#): Get text and email alerts about Restaurant Inspections, Potholes, 311 Service Requests, Vehicle Collisions, Street Lights, Consumer Complaints, Graffiti, Derelict Bikes, Sustainability.

[www.CoreData.nyc](#) New York City's housing and neighborhoods data hub, presented by the NYU Furman Center.

[NYC Community District Resources \(CUNY J School\)](#)

[Cops and Crime \(CUNY J School\)](#)

[Sentencing Project: State-level criminal justice data from a variety of sources](#)

[Census Microdata \(IPUMS\)](#)

[Federal data](#)

[Analytics USA: Federal website analytics](#)

[Beta.fec.gov + OpenFEC Campaign finance data improved.](#)

[OpenSecrets.org \(Political Donations in the US\)](#)

[US Extractive Industries Data](#)

[Foreign Assistance](#)

[Open Data Portals around the world](#)

[Data Is Plural](#): An awesome curated data selection by [Jeremy Singer-Vine](#), BuzzFeed data editor

[US Immigration Stats](#)

[FBI crime reports](#)

[NYC Education Data](#)

[College Statistics and Advanced Search](#)

[US Department of Education Data & Statistics](#)

[Bureau of Labor Statistics Data \(Inflation, Employment...\)](#)

[National Center for Education Statistics](#)

[Bureau of Justice Statistics](#)

[Bureau of Economic Analysis](#)

[Congressional Research Service reports \(option 1 and option 2\)](#)

[Environmental Protection Agency Data](#)

[Centers for Disease Control and Prevention Data \(Health Statistics\)](#)

[Gun Violence in the US](#)

[Dollars for Doctors](#)

[Surgeon Scorecard](#)

[Google Public Data Explorer](#)

[TRAC \(tracks and aggregates several types of government data\)](#)

GeoData

[Census Shapefiles](#) (cartographic boundary files in shapefile format for 2010 Census, 2000 Census and others). Includes counties, congressional districts, states, urban areas, zip code and more)

[CUNY KML Maps Database](#). Includes links to [KML files](#) in various categories (Arts & Culture, Buildings Business & Economics, Education, Environment, Government, Health & Science, Locations, Parks & Recreation, Politics, Public Safety, Transportation, US Census, ZIP Codes and Zoning)

Non-Profits/Charities

[Non Profits in the US](#): **Search the Nonprofit Network, by Center for Public Integrity.** “Nonprofits are not required to publicly disclose donors but must report money they give to other nonprofits on the public portion of their tax returns. By mining more than 850,000 tax forms filed electronically since

2010 from those givers, the Center for Public Integrity has peeled back a layer of secrecy surrounding the money fueling nonprofits. Follow the \$170 billion money trail with our Nonprofit Network tool.”
[Non Profits in the US: IRS 990 Filings on AWS](#). Machine-readable data from certain electronic 990 forms filed with the IRS from 2011 to present (2016) are available for anyone to use via Amazon S3. Form 990 is the form used by the United States Internal Revenue Service to gather financial information about nonprofit organizations. Data for each 990 filing is provided in an XML file that contains structured information that represents the main 990 form, any filed forms and schedules, and other control information describing how the document was filed. Some non-disclosable information is not included in the files.

[Charity Navigator](#)

[Non Profit Explorer](#), by ProPublica

[Guidestar 990 forms](#)

[Citizen Audit](#)

[Open Charities](#)

[Global Think Tank Directory](#) (includes annual Index)

[Data Worldbank](#)

[Data UN](#)

[Data IADB](#)

[Data World Health Organization](#)

[Data OECD](#)

[CIA World Factbook Country Comparisons](#)

[SIPRI Military and Defense \(Global Databases\)](#)

[Net Data Directory \(all things Internet, Infrastructure, IP, etc.\)](#)

[Global Social Progress Index](#)

[Enigma.io public global data explorer \(must signup\)](#)

[Data sets used in graphics published by Information is Beautiful](#)

[Gapminder Data - Global Indicators](#)

[Offshoreleaks and Panama Papers data](#)

[OpenCorporates \(Company Data\)](#)

[Investigative Dashboard Search \(Documents, People, Companies\)](#)

[The Observatory of Economic Complexity \(Country Reports, Imports, Exports, Trade, Products, Rankings\)](#)

[Let Me Get That Data For You: Enter a gov domain name and get filetype, url and title](#)

[List of states and territories of the United States](#)

[Americas Open Geocode \(AOG\) database](#): Includes Countries of the World, Country codes in English, Spanish, French; cities of the World, list of Government Websites for the Countries of the World, US City/ZipCodes, US State/Counties/Cities Dataset

[Find the Company](#)

[OpenCorporates](#)

[Graphik](#) (millions of data visualizations from available open data sports, business, politics, finance, gov. It also has data verticals)

Advanced Web Search

Beyond Google 2013, a curated list created by Margot Williams, from NPR for IRE June 2013

NSA: Untangling the Web http://www.nsa.gov/public_info/files/Untangling_the_Web.pdf

Find a person:

- AnyWho: <http://www.anywho.com>
- InfoBel: <http://www.infobel.com/en/world>
- Numberway <http://www.numberway.com/>
- Free lookups at MelissaData <http://www.melissadata.com/lookups/index.htm>

Find a person: Work

- LinkedIn <http://www.linkedin.com>
- CLEAR <http://www.clearhq.org/Directory>
- BRB verify professional licenses <http://verifyprolicense.com/>
- Salary database from PIBuzz <http://pibuzz.com/government-salaries/>
- Leadership directories \$\$ <http://www.leadershipdirectories.com/>

Search Public Records

- National Sex Offender Public Registry: <http://www.nsopw.gov>
- State Sex Offender Registry Web Sites: <http://www.fbi.gov/hq/cid/cac/states.htm>
- Directory of incarcerated offender databases at VineLink <http://www.vinelink.com>
- US federal inmate lookup <http://www.bop.gov/iloc2/LocateInmate.jsp>
- Finra BrokerCheck <http://www.finra.org/Investors/ToolsCalculators/BrokerCheck/>

- BRB Pubs Free Public Records <http://www.brbpub.com/free-public-records/>
- Search Systems <http://publicrecords.searchsystems.net/>
- Free Public Records Directory <http://publicrecords.onlinesearches.com/>
- CourtReference <http://www.courtreference.com/>
- NETR Online (real estate record directory) <http://www.netronline.com/>

Corporation Research

- ICIJ Offshore Leaks Database <http://offshoreleaks.icij.org/search>
- Investigative Dashboard (OCCRP) <http://www.investigativedashboard.org/>
- OpenCorporates.com <http://opencorporates.com>
- GlobalEdge Directory <http://globaledge.msu.edu/Global-Resources>
- Corporationwiki <http://www.corporationwiki.com/>
- Building database: Emporis <http://www.emporis.com/buildings>
- Commercial real estate: LoopNet <http://www.loopnet.com>

Non-Profits/Charities

- Guidestar US Form 990s <http://www.guidestar.org>
- Luke Rosiak's 990 database <http://lukerosiak.info/irs/>
- ProPublica Non-profit Explorer <http://projects.propublica.org/nonprofits/>
- OpenCharities.org UK <http://opencharities.org/>
- Charity Navigator <http://www.charitynavigator.org/>
- National Center for Charitable Statistics <http://nccsdataweb.urban.org/>
- Foundation Finder <http://foundationcenter.org/findfunders/foundfinder/>
- Global Think Tank Directory <http://www.gotothinktank.com/>

Watchlists, Wanted & Missing Persons Lists

- Corporate Intelligence Project <http://www.corpsearch.net/watchlists.html>
- Interpol Wanted Persons List <http://www.interpol.int/Wanted-Persons>
- Search OFAC <http://sdnsearch.ofac.treas.gov/>
- SAM.gov: Debarred from business with US <http://www.sam.gov>
- Denied Persons List http://export.gov/ecr/eg_main_023148.asp
- ICRC: missing in Somalia <http://www.familylinks.icrc.org/eng/missing-somalia>

Data and Statistics

- Data.gov <http://catalog.data.gov/dataset>
- Official Statistics on the Web <http://www.offstats.auckland.ac.nz>
- Population Reference Bureau DataFinder <http://prb.org/datafinder.aspx>
- Guardian DataStore World Government Data
<http://www.guardian.co.uk/world-government-data>
- US Census Bureau International Database<http://www.census.gov/population/international/data/idb/informationGateway.php/>

International planes & ships finders

- Pirates! NATO Daily Piracy Update
<http://www.shipping.nato.int/operations/OS/Pages/DailyPiracyUpdate.aspx>
- Equasis <http://www.equasis.org>
- MarineTraffic.com <http://www.marinetraffic.com/ais/>
- Ship to shore phone directory: <http://www.inmarsat.com/shipsdirectory/index.htm>
- FlightAware: <http://www.flightaware.com>
- Flightwise <http://flightwise.com/>
- Pinkfroot <http://my.pinkfroot.com/> Planefinder, Shipfinder, iPhone apps
- FAA Aircraft registryhttp://www.faa.gov/licenses_certificates/aircraft_certification/aircraft_registry/interactive_aircraft_inquiry/
- Landings: <http://www.landings.com>
- Airport Codes: <http://www.airlinecodes.co.uk/aptcodesearch.asp>
- US work visa data http://www.foreignlaborcert.doleta.gov/performance_data.cfm
- TRACE Compendium<https://secure.traceinternational.org/Knowledge/Compendium.html>

Interesting Databases

- US work visa data http://www.foreignlaborcert.doleta.gov/performance_data.cfm
- TRACE Compendium<https://secure.traceinternational.org/Knowledge/Compendium.html>
- SpyPedia \$\$
- Guantanamo Docket (New York Times/NPR)<http://projects.nytimes.com/guantanamo>

FOIAs, leaks and deletes

- Government Attic <http://www.governmentattic.org/>
- JunketSleuth “exposing government travel excesses” <http://junketsleuth.com/>
- Public Intelligence <http://publicintelligence.net/>
- Archive-it <http://archive-it.org/explore?show=Collections>
- netpreserve.org <http://netpreserve.org/about/archiveList.php>
- CyberCemetery defunct agencies <http://govinfo.library.unt.edu/>
- Internet Archive Wayback Machine <http://www.archive.org>

Lobbying for US and non-US interests

Senate Office of Public Records Lobbying Disclosure Act

http://www.senate.gov/legislative/Public_Disclosure/LDA_reports.htm

US State Dept. Foreign Agents Registration Act (FARA) since 1940s

<http://www.usdoj.gov/criminal/fara/>

US Lawmakers & Officials Disclosure

Ethics.data.gov <http://ethics.data.gov>

Legistorm <http://www.legistorm.com>

Office of Government Ethics

<http://www.oge.gov/Open-Government/Access-Records/Current-Executive-Branch-Nominations-and-Appointments/>

Center for Responsive Politics <http://www.opensecrets.org/pfds/>

Campaign Finance FEC <http://www.fec.gov/disclosure.shtml>

National Archives

- Archives I and Archives II and Presidential Libraries www.archives.gov
- Online Public Access <http://www.archives.gov/research/search/index.html>
- Archival Databases <http://aad.archives.gov/aad/>

Digital library access

- JSTOR Register & Read: Early Journal Content <http://www.jstor.org/betasearch>
- Digital Public Library of America <http://dp.la/>
- Hathi Trust Digital Library <http://www.hathitrust.org>
- Internet Archive <http://archive.org/index.php>
- Library card: NYPL, DC Public Library, etc

Web Research Updates

- Toddington International <http://www.toddington.com/resources/>
- Gary Price InfoDocket <http://infodocket.com/>
- FullTextReports <http://fulltextreports.com/>
- PI Buzz (Tamara Thompson) <http://pibuzz.com/>

(this list was compiled by Knight Center for Journalism in the Americas)