Some Cool Datasets

*Compiled by Lindsey, but made better by you. Please add more data!*

**People:**

* [Census Factfinder](https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml): Information on people, places, economies, businesses, education, demographics, etc by U.S., state or county. Contains multiple years with more accessible in other parts of the Census site. Includes easy export to spreadsheets as well as visualization tools.
  + Start with the guided search if you don’t know where to begin.
  + [Tutorials here](https://www.census.gov/library/video/2017/tutorial-excel-census-data.html).
* [Census Reporter](https://censusreporter.org/): Prettier and easier to use, but not as much information. Great for backgrounding a place.
* [IPUMS](https://www.ipums.org/): Microdata from the Census as well as other countries
* [General Social Survey](http://gss.norc.org/): Tracks opinions and demographics in the U.S. since 1972.
* [Bureau of Labor Statistics](https://www.bls.gov/data/): Information on prices, inflation, employment, different professions, pay and benefits, productivity, etc.
  + American Time Use Survey
  + “Economy at a glance” for states and large cities. [Sample](https://www.bls.gov/eag/eag.ga.htm)
* [Child Trends](https://www.childtrends.org/databank-indicators/databank-by-topics): Aggregates hundreds of statistics on children and families across many topics.

**Government:**

* [Data.gov](https://www.data.gov/): “The home of the U.S. Government’s open data.”
* DOJ’s [Foreign Agents Registration database](https://www.data.gov/) ([document search](https://efile.fara.gov/pls/apex/f?p=145:10))
  + More justice data is housed on the [Bureau of Justice Statistics site](https://www.bjs.gov/)
* [Plum Book](https://www.govinfo.gov/collection/plum-book?path=/gpo/United%20States%20Government%20Policy%20and%20Supporting%20Positions%20(Plum%20Book)/2016): Lists employees and salaries in the executive branch
* [Congressional Misconduct Database](https://www.govtrack.us/misconduct) (Gov Track)
* Bureau of Transportation Statistics: [on-time airline departure](https://www.transtats.bts.gov/Tables.asp?DB_ID=120) and hundreds of other metrics
* IRS 990 Filings: For tracking charities.
  + Individual charities: [ProPublica’s Nonprofit Explorer](https://projects.propublica.org/nonprofits/) or [Charity Navigator](https://www.charitynavigator.org/)
  + Industry trends: [Amazon](https://registry.opendata.aws/irs990/) (contact Jeremy/Rachel)
* [Office of Management and Budget](https://www.whitehouse.gov/omb/historical-tables/): Lots of great data on the federal budget. There’s also a list of OMB’s data inventory [here](https://obamawhitehouse.archives.gov/open/around/eop/omb/datasets).
* [The Congressional Budget Office](https://www.cbo.gov/about/products/budget-economic-data): Data on historical budgets and projections that accompany Congress reports
* [Federal contract misconduct](https://www.contractormisconduct.org/misconduct?misconduct_type=&disposition_type=&court_type=&enforcement_agency=&contracting_party=&penalty=100000000-999999999999&from=2014&to=2019) database

**Crime and Justice:**

* [FBI Uniform Crime Reporting](https://www.fbi.gov/services/cjis/ucr)**:** National and more local data on crimes, use of force, hate crimes and police
* [Bureau of Justice Statistics](https://www.bjs.gov/index.cfm?ty=dca): Data on justice spanning multiple types of federal, state and local prison and jails, including parole, victims and crime. Has good historical data.

**NY-Specific Data:**

* [SeeThroughNY](https://www.seethroughny.net/): Database of public employees and public spending compiled by a non-partisan think tank.
* [New York State Crime Data](http://www.criminaljustice.ny.gov/crimnet/ojsa/stats.htm): State level data on all sorts of crimes, including arrest and disposition data.
* [NYPD Supplementary Homicide Reports](https://www1.nyc.gov/site/nypd/stats/reports-analysis/homicide.page): Data on when, where, why and how murders occur.
* [NYC Open data portal](https://opendata.cityofnewyork.us/data/)
* [NY State open data portal](https://data.ny.gov/)

**Economics and Business:**

* [S&P Capital IQ](https://www.capitaliq.com/ciqdotnet/login-sso.aspx?bmctx=202C7940E5B3C784A1D977EF9E24AB0D&contextType=external&username=string&enablePersistentLogin=true&OverrideRetryLimit=0&contextValue=%2Foam&password=secure_string&challenge_url=https%3A%2F%2Fwww.capitaliq.com%2Fciqdotnet%2Flogin-sso.aspx&request_id=7309555740275945944&authn_try_count=0&locale=en_US&resource_url=https%253A%252F%252Fwww.capitaliq.com%252FCIQDotNet%252FLogin.aspx): Wealth of information on publicly traded companies pulled from SEC filings.
* [Bureau of Labor Statistics](https://www.bls.gov/Data/): Data on pay and benefits, industries, employment and unemployment, inflation, how people use their time, productivity and economic profiles of local areas.
* [Bureau of Economic Analysis](https://www.bea.gov/data): Easy to use site that has a wealth of economic data
* [Federal Reserve Bank of St. Louis’s FRED](https://fred.stlouisfed.org/)

**War:**

* [Long War Journal](https://www.longwarjournal.org/us-airstrikes-in-the-long-war): Airstrike data and Taliban control in Afghanistan
* [Defense Logistics Agency](http://www.dla.mil/DispositionServices/Offers/Reutilization/LawEnforcement/PublicInformation.aspx): Military weapons being sold to local police departments
* [National Consortium for the Study of Terrorism and Responses to Terrorism](http://www.start.umd.edu/)
* [Stockholm International Peace Research Institute](https://www.sipri.org/databases): Data on military and arms spending, trade, exports and embargoes internationally

**Transportation:**

* [National Highway Traffic Safety Administration](https://www.nhtsa.gov/research-data): driver registration, crashes, traffic records, recalls
  + [Fatality Analysis Reporting System](https://www.nhtsa.gov/research-data/fatality-analysis-reporting-system-fars): Comprehensive information on factors in car crashes including information on driver, car, place, weather conditions, etc

**Health:**

* [CDC](https://www.cdc.gov/datastatistics/index.html): Health data including births, deaths, diseases, obesity, smoking, life expectancy, including demographic information. Some surveys too like the [youth risk behavior survey](https://www.cdc.gov/healthyyouth/data/yrbs/index.htm).
* [Centers for Medicare & Medicaid Services](https://www.cms.gov/openpayments/): Payment data showing procedures and cost, including names for doctors. Large databases [here](https://www.cdc.gov/nchs/data_access/ftp_data.htm).
* [EPA](https://www.epa.gov/outdoor-air-quality-data): Air quality data at outdoor monitors across the U.S.
* [Consumer Product Safety Commission](https://www.cpsc.gov/Newsroom/Downloadable-Data):
* [HCUP](https://www.hcup-us.ahrq.gov/db/nation/nrd/nrddbdocumentation.jsp): Longterm outcomes for patients in U.S. hospitals
* [US News & World Report Hospital Rankings](https://www.usnews.com/static/documents/health/best-hospitals/BH_Methodology_2018-19.pdf): They have a lot of data and you can look at the methodology to see where it all comes from.
* [County Health Rankings](http://www.countyhealthrankings.org/): Aggregates a lot of data on the county level.
* [Kaiser Family Foundation](https://www.kff.org/statedata/): Lots of data on health and health disparities

**Education:**

* [Department of Ed](https://www2.ed.gov/rschstat/catalog/index.html): List of all the databases they have from campus safety, to state education spending, to equity in school athletics, including those that are housed on other websites.
* [National Center for Education Stats](https://nces.ed.gov/): Easy to use data, mostly at national and state with more high-level education breakdowns for K-12 and universities such as graduation rates, tuition, etc.
* [Civil Rights Collection](https://ocrdata.ed.gov/): Super rich data with very specific breakdowns. Enrollment by race and disability status for school, district, state and national level, including enrollment, access to advanced courses, chronic absenteeism, teacher resources and certification and discipline.

**Environment**

* [Toxics Release Inventory (TRI) Program](https://www.epa.gov/toxics-release-inventory-tri-program/tri-data-and-tools): Data on toxic chemicals managed as waste released into the environment.
* [CO2 emissions (metric tons per capita)](https://data.worldbank.org/indicator/en.atm.co2e.pc): World Bank database of CO2 emissions per country per capita
* [Outdoor Air Quality Data](https://www.epa.gov/outdoor-air-quality-data/download-daily-data): Outdoor daily pollution data based on pollution monitors.
* [NOAA Temperature data](https://www.ncdc.noaa.gov/cag/city/time-series/USW00023185/tavg/1/2/1895-2018?base_prd=true&firstbaseyear=1901&lastbaseyear=2000&trend=true&trend_base=10&firsttrendyear=1895&lasttrendyear=2018): Temperature data for a given city/region dating back to the 1800s.
* [Energy Information Administration](https://www.eia.gov/): Data on energy sources for U.S. and other countries.

**Immigration:**

* [Department of Homeland Security Yearbooks](https://www.dhs.gov/immigration-statistics/yearbook): Annual breakdowns of legal and unauthorized immigration.
* [Customs and Border Protection](https://www.cbp.gov/newsroom/media-resources/stats): Data on border resources and staffing.
* [US Customs and Immigration Services](https://www.uscis.gov/tools/reports-studies/immigration-forms-data): Data by entry program.

**International:**

* [WTO](http://stat.wto.org/Home/WSDBHome.aspx): Data on international trades and tariffs.
* [World Bank](https://data.worldbank.org/): TONS of data spanning beat areas, available by country, but also on a micro level.
* [UN](http://data.un.org/): Data spanning beat areas on different countries.
* [Eurostats](https://ec.europa.eu/eurostat/data/database): Collects data across government agencies for European countries.
* [CIA Factbook](https://www.cia.gov/library/publications/the-world-factbook/): Data-driven profiles of most countries.
* [Foreign Assistance](https://www.foreignassistance.gov/explore): Data sets/interactive maps of U.S. foreign assistance.
* [OECD](https://stats.oecd.org/): Econ and social stats from around the industrialized world.

**Sports:**

* [Baseball Reference](https://www.baseball-reference.com/), [Pro Football Reference](https://www.pro-football-reference.com/)**,** [Basketball Reference](https://www.basketball-reference.com/)**:** packed with data on just about any game in history as well as individual player stats. There are also hockey, soccer and college basketball and football sites.

**Grab Bag**

* [National Vulnerability Database](https://nvd.nist.gov/) for technology vulnerabilities.
* [Coin metrics](https://coinmetrics.io/data-downloads/): Trends for different types of new currency.
* [Survey of Consumer Finances](https://www.federalreserve.gov/econres/scfindex.htm?elqCampaignId=5888)
* [Google Trends](https://trends.google.com/trends/?geo=US)
* [Google Dataset Search](https://toolbox.google.com/datasetsearch) like Google search, but for datasets.
* A list of [all the datasets](https://data.world/jsvine/data-is-plural-archive/workspace/file?filename=Data+Is+Plural+%E2%80%94+Structured+Archive.xlsx) published in the Data Is Plural newsletter (you have to sign up for a free data.world account to view this).