storytelling with data®

where to find data: an ever-growing list of public sources

Here are some places to find/access publicly available data. Help us make this list comprehensive: if you know of other sources we should include, please email us.

- **American Community Survey** - the premier source for detailed information about the American people and workforce.
- **American Institute of Physics** - online oral history transcripts
- **American Soccer Analysis** - all soccer stats for MLS (the soccer league in the USA and Canada)
- **ArcGIS Hub Open Data** - Localized geo-spatial data.
- **Armed Conflict Location & Event Data Project (ACLED)** - records global data on political violence and protest events
- **The Association of Religious Data Archives** - A collection of surveys, polls, and other data submitted by researchers and made available online by the ARDA.
- **Awesome JSON Datasets** - GitHub repository with links to data on bitcoin, cars, climate, and more.
- **Awesome Public Datasets** - GitHub repository, topic-centric list of high-quality open datasets in public domains. By everyone, for everyone!
- **BJS.gov** - Bureau of Justice statistics and data.
- **BRFSS** - Behavioral Risk Factor Surveillance System
- **British Oceanographic Data Center** - Publicly accessible marine data.
- **Bureau of Transportation Statistics** - Preeminent source of statistics on commercial aviation, multimodal freight activity, and transportation economics; provides context to decision makers and the public for understanding statistics on transportation.
- **CDC** - Center for Disease Control datasets by topic.
- **CDC Wonder Online Database** - WONDER online databases utilize a rich ad-hoc query system for the analysis of public health data.
- **Canada** - open data for the federal government of Canada
- **Child Trends DataBank** - Child Trends improves the lives and prospects of children and youth by conducting high-quality research and sharing the resulting knowledge with practitioners and policymakers.
- **City of Chicago** - open data portal for all things Windy City
- **COVID-19 School Data Hub** - A central database for educators, researchers, and policymakers to understand how the COVID-19 pandemic shaped US students' modes of learning in 2020-21.
- **Data.gov** - The home of the U.S. Government's open data.
- **Data.gov.au** - Australian government's open data.
- **Data.gov.in** - Indian government's open data.
- **Data.gov.nz** - New Zealand government's open data.
- **Data.gov.uk** - Data published by UK government departments and agencies, public bodies, local authorities.
- **Data.gov.sg** - Singapore government's open data.
- **Data is Plural** - Spreadsheet of public datasets and sources.
- **DataSF** - Search hundreds of datasets from the City and County of San Francisco.
- **Dataverse** - Open source research data repository software.
- **Data.world** - Data aggregator of public datasets, projects and sources.
- **Diversity Data Kids** - Data on diversity, opportunity and equity amongst US children
- **DRYAD** - Curated general-purpose repository that makes the data underlying scientific publications discoverable, freely reusable, and citable.
- **The Economist** - markets & data.

- **EIA** - US Energy Information Administration's energy data.
- **Enigma Public** - "The world's broadest collection of public data".
- **Environmental Information Data Centre -** a NERC Data Centre hosted by the UK Centre for Ecology & Hydrology (UKCEH) concerned with the terrestrial and freshwater sciences.
- **Eurocontrol** - the European equivalent of the FAA, publishes traffic and CO2 emissions data and archive
- **Eurostat** - European statistics.
- **FAA Data** - Data and statistics from Federal Aviation Administration research.
- **FAOSTAT Data** - Food and Agriculture Organization of United Nations datasets.
- **FBI Crime Data Explorer** - View trends, download bulk datasets, and access the Crime Data API for reported crime at the national, state, and agency levels.
- **FEMA Data Feeds** - FEMA data available in multiple formats.
- **Figshare** - Opening up scientific data to the world.
- **FiveThirtyEight** - Datasets used in its articles available via Github repository.
- **FlowingData** - 30 resources to find the data you need.
- **Gapminder.org** - Data in Gapminder World (over 500 entries).
- **The General Social Survey** - Datasets from NORC.org on what Americans think and feel about such issues as national spending priorities, crime/punishment, intergroup relations, and confidence in institutions.
- **Google Public Data Explorer** - Google's Public Data Explorer provides public data and forecasts from a range of international organizations and academic institutions.
- **Health Intelligence** - Links to a ton of health related datasets.
- **Healthcare Cost and Utilization Project** - Free healthcare statistics.
- **HealthData.gov** - dedicated to making high value health data more accessible to entrepreneurs, researchers, and policy makers in the hopes of better health outcomes for all.
- **Health and Retirement Study** - The University of Michigan Health and Retirement Study (HRS) is a longitudinal panel study that surveys a representative sample of approximately 20,000 people in America, supported by the National Institute on Aging (NIA U01AG009740) and the Social Security Administration
- **Human Fertility Database** - Detailed and high-quality data on period and cohort fertility to facilitate research on changes and inter-country differences in fertility in the past and in the modern era.
- **Human Mortality Database** - Provides detailed mortality and population data to researchers, students, journalists, policy analysts, and others interested in the history of human longevity.
- **The Huntington's Early California Population Project** - Public access to all the information contained in California's historic mission registers.
- **IIE.org Open Doors Data Portal** - Data tables for the past decade on international students and scholars in the United States and on American students studying abroad for academic credit.
- **Informatics Lab Open Data** - 80TB of meteorological data.
- **IPUMS** - Provides census and survey data from around the world integrated across time and space.
- **Kaggle** - Open datasets on variety of topics (government, health, science, popular games, dating trends).
- **KIDS COUNT Data Center** - Provides data on children and families
- **KDNuggets Datasets** - Lists a ton of data repositories.
- **Lahman's Baseball Database** - Complete batting and pitching statistics from 1871 to 2016, plus fielding statistics, standings, team stats, managerial records, post-season data, and more.
- **London Datastore** - The London Datastore is a free and open data-sharing portal where anyone can access data relating to London.
- **Makeover Monday** - Datasets from the weekly Makeover Monday competition.
- **MEPS** - Medical Expenditure Panel Survey
- **National Center for Education Statistics** - primary federal entity for data related to education.
- **National Transit Database** - Congress established the NTD to be the Nation's primary source for information and statistics on the transit systems of the United States.
- **NASA** - National Aeronautics and Space Administration datasets and images.

- **NASA Earth Data** - Data from The Earth Observing System Data and Information System (EOSDIS).
- **NHAMCS** - National Hospital Ambulatory Medical Care Survey Data.
- **NHANES** - National Health and Nutrition Examination Survey
- **New York City** - Access information that is produced and used by New York City government.
- **New York State** - Browse and download over 1,500 New York State data resources on topics ranging from farmers' markets to solar photovoltaic projects to MTA turnstile usage.
- **ODI Leeds** - Pioneer nodes bring together communities, host events, produce open data projects and help promote understanding of open data in our region and worldwide.
- **OECD Data** - Find, compare and share the latest OECD data: charts, maps, tables and related publications.
- **Office for National Statistics** - The UK's largest independent producer of official statistics and the recognised national statistical institute of the UK.
- **Toronto Open Data** - Open data from the provincial government of Ontario, Canada
- **Open Data Nederland** - Open data from the Netherlands.
- **Open Data Network** - Publish data and share. Find data and build. Answer questions.
- **Opportunity Atlas** - data on childhood outcomes and inequality in the US.
- **Our World in Data** - Explore the ongoing history of human civilization at the broadest level, through research and data visualization (with downloadable datasets).
- **PermID from Thomson Reuters** - PermIDs are open, permanent and universal identifiers where underlying attributes capture the context of the identity they each represent.
- **Pew Research Center** - Download datasets.
- **Plenar.io** - centralized hub for open datasets from around the world
- **PolicyMap** - Easy-to-use online mapping with data on demographics, real estate, health, jobs and more in communities across the US.
- **Princeton University Library** - Political opinion & survey data sources.
- **re3data.org** - Registry of research data repositories.
- **Retrosheet.org** - Baseball play by play and game log stats back to the 20s (see "data downloads").
- **Satori** - open data initiative
- **Statista** - Statistics portal for stats & studies from 18,000 sources.
- **St. Louis Federal Reserve** - Download 507,000 US & international time series from 87 sources.
- **SAMHSA** - Substance Abuse and Mental Health Services Administration.
- **Swiss Open Government Data** - opendata.swiss is the portal for Swiss open government data (OGD).
- **Southern Oral History Program** - digital transcripts and audio from interviews
- **Tableau Public** - Datasets on education, government, science, lifestyle and more.
- **Tidy Tuesday** - dataset repository for weekly data projects aimed at the R ecosystem
- **Toronto Open Data** - Open data catalogue from city of Toronto.
- **Transport for London** - TFL data feeds.
- **UC Irvine Machine Learning Repository** - Nearly 400 datasets maintained as a service to the machine learning community. You may view all data sets through our searchable interface.
- **UMich ICPSR** - University of Michigan's searchable database.
- **UN Comtrade Database** - Trade statistics from the United Nations.
- **Uniform Crime Reporting** - FBI datasets.
- **University of Michican Research Guides** - data repository for qualitative datasets
- **University of Rochester LibGuide "Data-Stats"** - Repository listing many public data sources & datasets.
- **US Bureau of Labor Statistics** - National longitudinal surveys. See also "Databases & Tables by Subject."
- **US Census Bureau Data** - Browse by topic, download data.
- **USA Facts** - Federal, state and local data from over 70 government sources
- **Utah Open Data** - State of Utah open data catalog with data provided by State of Utah agencies.
- **Varieties of Democracy** - Time series data on various aspects of democracy.
- **Western Pennsylvania Regional Data Center** - Now serving 262 datasets and counting.

- **WHO Data Repository** - World Health Organization database.
- **The World Bank Data Catalog** - 20K+ datasets, this catalog lists all open data and other publicly available data disseminated through World Bank sites.
- **World Values Survey** - Data from nationally representative surveys conducted in almost 100 countries using a common questionnaire to measure changing values and impact on social and political life.
- **World Wealth & Income Database** - Download open access datasets.

---

storytellingwithdata.com