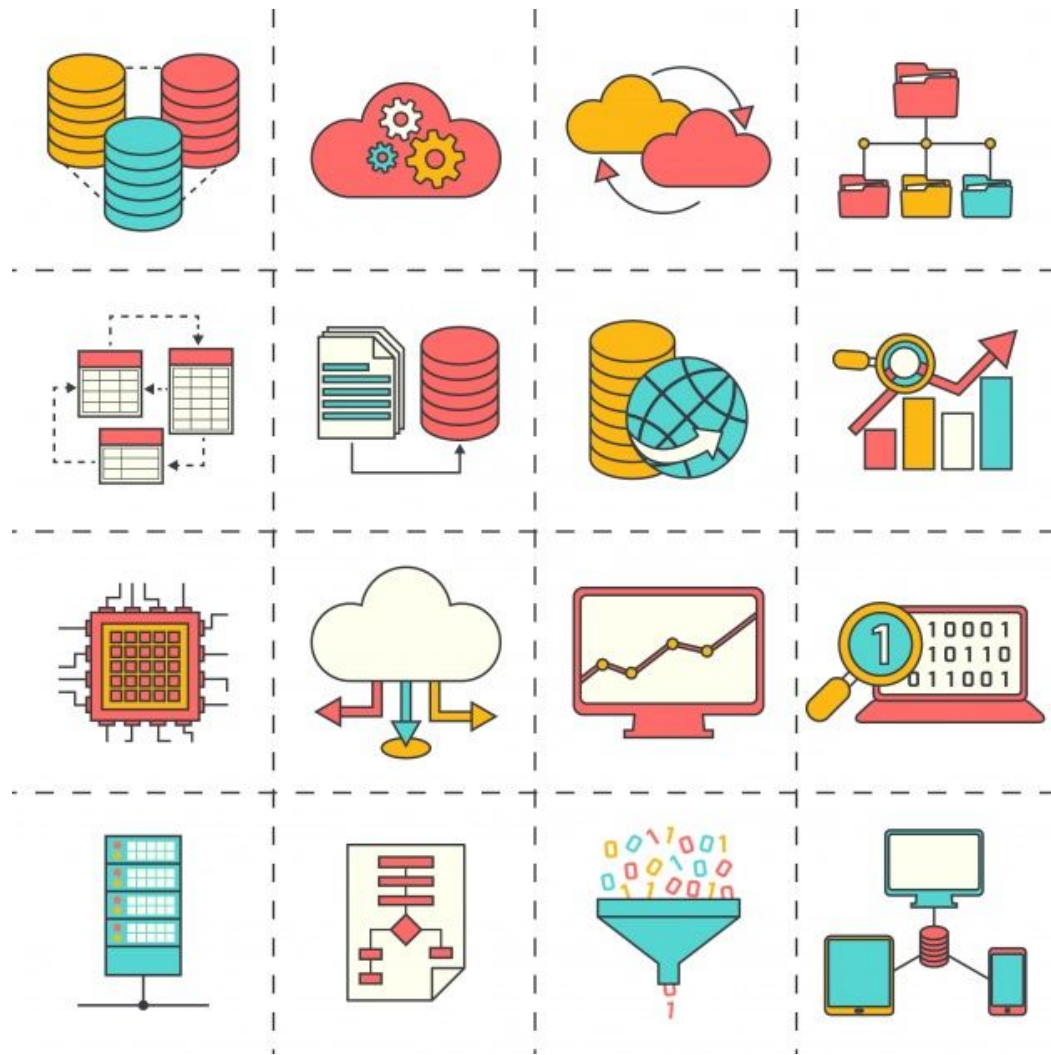


What is data and where to find it



Getting and Cleaning Data



Tab delimited

Name	Sex	Weight	Height
Alfred	M	75	163
Carl	M	80	180
Eleanor	F	72	175
Georgia	F	68	172
Ian	M	77	168
Karen	F	76	173
Mike	M	80	190

Comma separated

Name,Sex,Weight,Height
Alfred,M,75,163
Carl,M,80,180
Eleanor,F,72,175
Georgia,F,68,172
Ian,M,77,168
Karen,F,76,173
Mike,M,80,190



Website with information
you want

Structured

Unstructured

APIs

Scraping

Structured, usable
dataset



Restaurants

ID	Name	Address	Type
39	Jenny's Pizza	123 Main St	Pizza
128	Park Ave Pasta	999 Park Ave	Pasta
47	Raw	1 East Ave	Sushi
29	Noodles	12 Western Blvd	Ramen
482	Sakura	700 Broadway St	Japanese
163	Wakey, wakey	122 Monument Ave	Breakfast

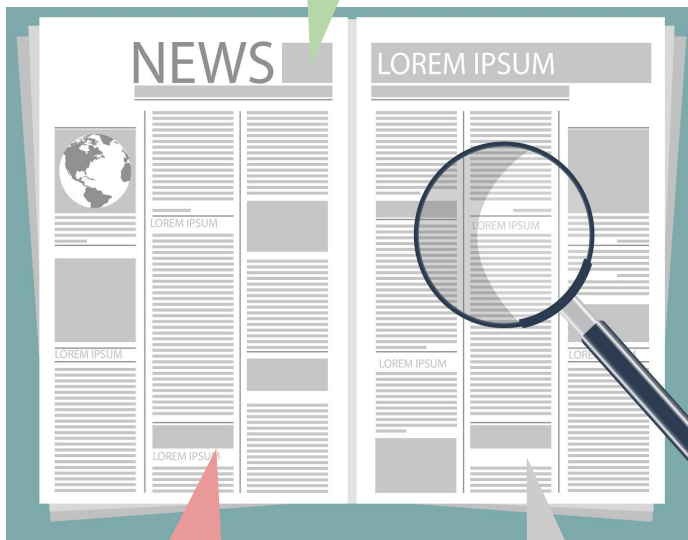
Inspections

Restaurant	Score	Inspector
39	10	2
39	9	3
482	4	5
47	8	4
163	7.5	2
47	1	2

Inspectors

ID	Name
1	Jonathan Blue
2	Henry Green
3	Rebecca Pink
4	Paul Brown
5	Jennifer Yellow

Positive:
70%



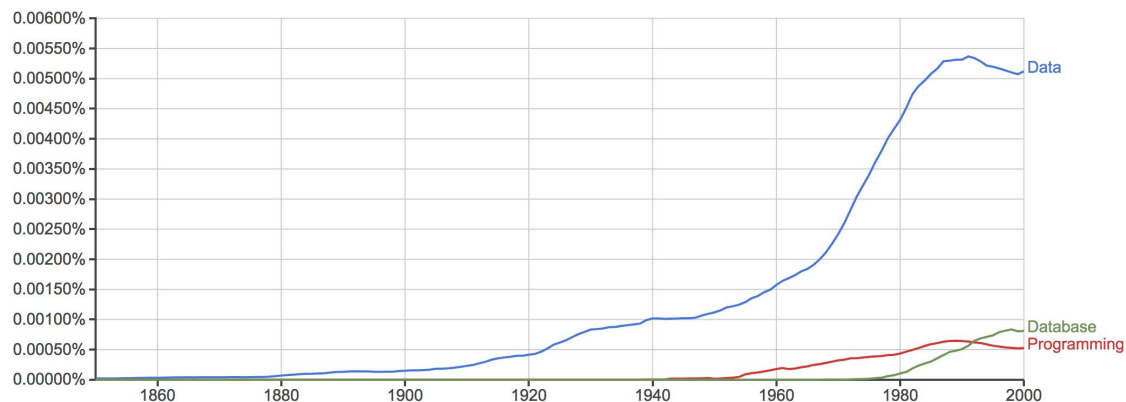
Negative:
20%

Neutral:
10%

Google Books Ngram Viewer

Graph these comma-separated phrases: ☐ case-insensitive

between and from the corpus with smoothing of [Search lots of books](#)



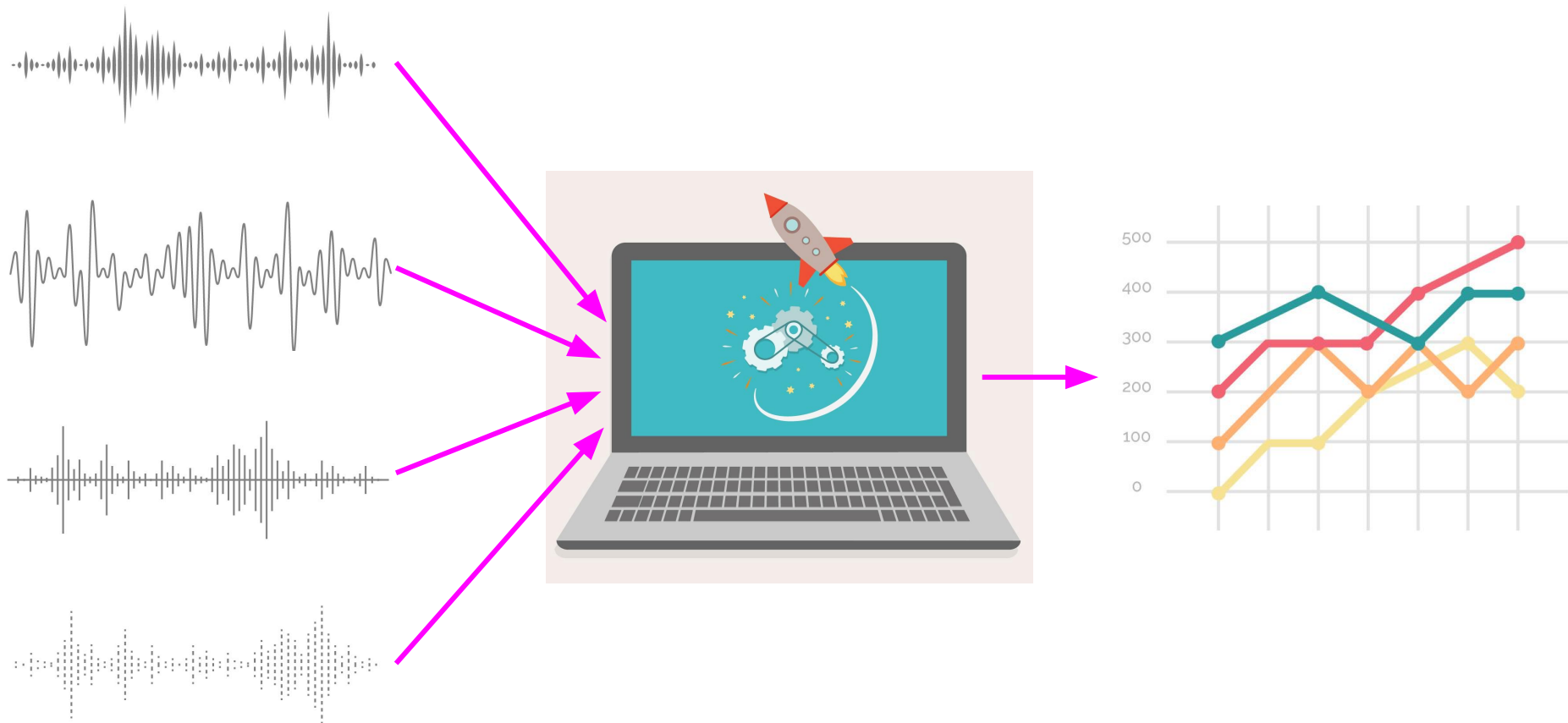


The cars in this image of a Brooklyn neighborhood can reveal a lot about the residents there. // Google Street View

Google Street View Can Reveal How Your Neighborhood Votes

LINDA POON DEC 6, 2017





Private vs. Public Data

PRIVATE DATA

- Access is limited
- May contain sensitive information



Created by Zarubin-leonid - Freepik.com

PUBLIC DATA

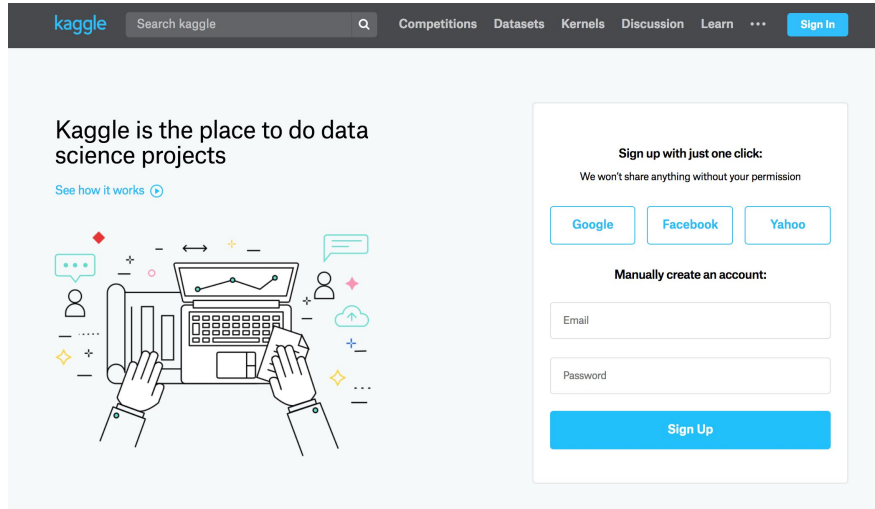
- Open Access
- Freely-available



Created by Kjpargeter - Freepik.com

Open Data Repositories

www.kaggle.com



The image shows the Kaggle homepage. At the top is a dark navigation bar with the Kaggle logo, a search bar, and links to Competitions, Datasets, Kernels, Discussion, Learn, and a Sign In button. The main content area has a light blue background. On the left, there's a section titled 'Kaggle is the place to do data science projects' with a link 'See how it works'. Below this is an illustration of hands interacting with a laptop displaying a line graph, surrounded by various data-related icons. On the right, there's a sign-up form. It starts with the heading 'Sign up with just one click:' and a subtext 'We won't share anything without your permission'. Below this are three buttons for Google, Facebook, and Yahoo. Then, it says 'Manually create an account:' followed by input fields for Email and Password, and a large blue 'Sign Up' button at the bottom.

Kaggle is the place to do data science projects

See how it works

Sign up with just one click:
We won't share anything without your permission

Google Facebook Yahoo

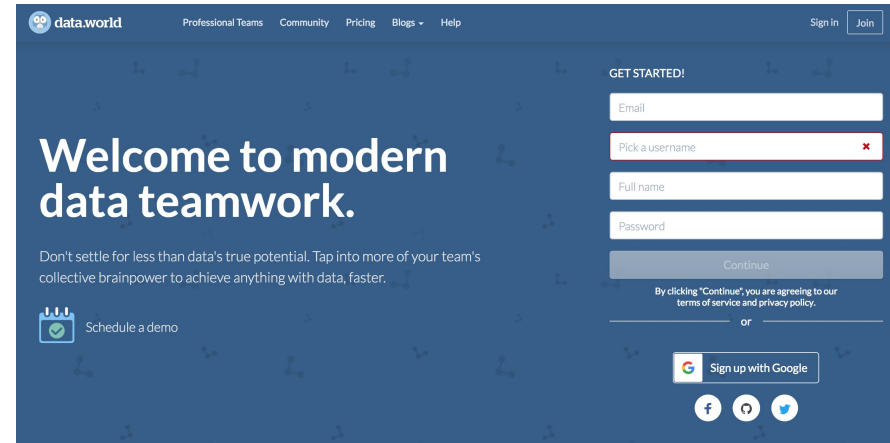
Manually create an account:

Email

Password

Sign Up

data.world



The image shows the data.world homepage. It has a dark blue background with a pattern of small white icons. The top navigation bar includes the data.world logo, links to Professional Teams, Community, Pricing, Blogs, and Help, and Sign In/Join buttons. The main content area features a large heading 'Welcome to modern data teamwork.' followed by a paragraph: 'Don't settle for less than data's true potential. Tap into more of your team's collective brainpower to achieve anything with data, faster.' Below this is a 'Schedule a demo' button. On the right side, there's a 'GET STARTED!' section with input fields for Email, Pick a username (with a red 'x' icon), Full name, and Password, followed by a 'Continue' button. Below the 'Continue' button is a small disclaimer: 'By clicking "Continue", you are agreeing to our terms of service and privacy policy.' At the bottom right, there's a 'Sign up with Google' button and social media icons for Facebook, GitHub, and Twitter.

data.world Professional Teams Community Pricing Blogs Help Sign In Join

GET STARTED!

Email

Pick a username

Full name

Password

Continue

By clicking "Continue", you are agreeing to our terms of service and privacy policy.

Schedule a demo

Sign up with Google

f g t

Search datasets...



Order by:

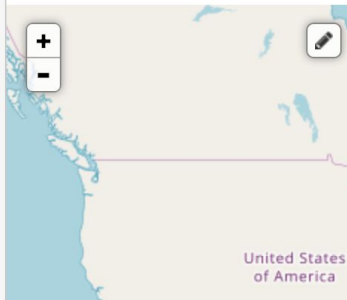
Popular ▾

Datasets ordered by Popular

Filter by location

Clear

Enter location...

Map tiles & Data by [OpenStreetMap](#), under [CC BY SA](#).

Topics



Clear All

285,732 datasets found

ZIP Code Data  1047 recent views

Department of the Treasury — This study provides detailed tabulations of individual income tax return data at the state and ZIP code level.

HTML

Demographic Statistics By Zip Code  977 recent views

City of New York — Demographic statistics broken down by zip code

CSV RDF JSON XML

U.S. Hourly Precipitation Data  699 recent views

National Oceanic and Atmospheric Administration, Department of Commerce — Hourly Precipitation Data (HPD) is digital data set DSI-3240, archived at the National Climatic Data Center (NCDC). The

Federal

City

Federal

Midwest Home to Most of the Counties With Decreases in Median Age

Approximately half (51.4 percent) of the nation's 531 counties that were getting younger between April 2010 and July 2017 were in the Midwest, according to newly released 2017 population estimates.



POPULATION CLOCK



U.S. Population

328,116,285

World Population

7,485,525,421

Jul 12, 2018 17:09 UTC (+4)

[Learn More >>](#)

QUICKFACTS



Did You Know

30.2%
of persons 25 years+ in Searcy city, Arkansas
have a bachelor's degree or higher

Source: American Community Survey

Select a state to begin



U.S. Census Bureau Economic Indicators



	Monthly Wholesale Inventories	\$633.5 B	↑
	May 2018 Report		0.6%
	Released 10:00 AM EDT, 7/11/18		
	International Trade: Goods & Services	\$43.1 B	↓
	May 2018 Report	Deficit	6.6%
	Released 8:30 AM EDT, 7/6/18		
	Manufacturers' Goods	\$498.2 B	↑
	View All		
		* change not statistically significant	
		° significance not reported / applicable	

Latest News



Schedule for Income, Poverty and Health Insurance Statistics and ACS

July 09, 2018

Schedule for the release of the 2017 income, poverty and health insurance coverage statistics and the 2017 American Community Survey.



U.S. CITY OPEN DATA CENSUS

POWERED BY OPEN DATA CENSUS

[About](#) [Changes](#) [FAQ](#) [Tutorial](#)Number of places **264**Number of datasets **679**Number of open datasets **381**Percentage open **56%**

Welcome to the U.S. City Open Data Census! Submissions for the 2018 edition of the census [are now open](#)! See [past years' submissions here](#).

[Read more](#) about what has changed from the old version. At the moment, cities are ordered alphabetically; you can order by rank by double-clicking the "Rank" header.

Search

Share this page

[Twitter](#)[Facebook](#)[G+ Google+](#)

Rank	Place	Budget	Crime Reports	Parcels	Construction Permits	Zoning	Service Requests	Code Violations	Employee Salaries	Business Listings	Spending	Restaurant Inspections	Public Facilities	Traffic Crashes	Property Assessment	Procurement Contracts	Emergency Calls	Lobbyist Activity	Use-of-Property	Property Transfers	Website Analytics	Score
86	Abilene, TX	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	0%
50	Alameda County, CA	<div></div>	<div></div>	<div></div>	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	<div></div>	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	+ Add	18%

TRACKING THE STATE OF OPEN GOVERNMENT DATA

The Global Open Data Index provides the most comprehensive snapshot available of the state of open government data publication



Compare countries

Ranked table and map views of participating countries.



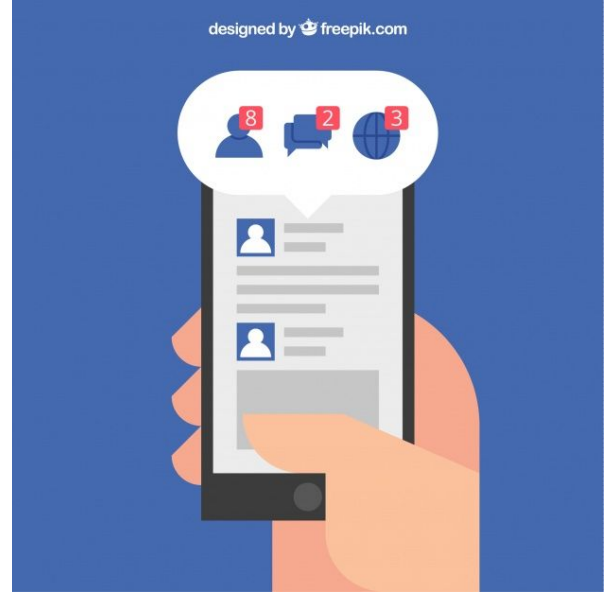
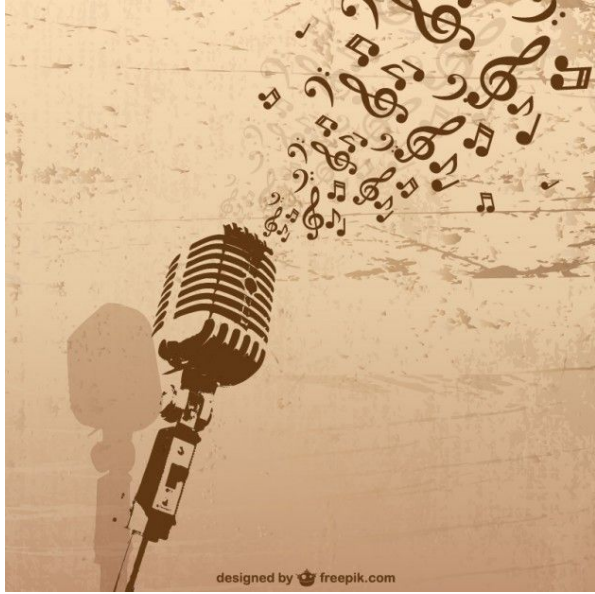
Discuss findings

Discuss your findings in our forum.



Get the insights

Insights from local and thematic perspectives.



Summarizing: What is data and where to find it



Getting and Cleaning Data