



Freely Available Datasets

UC Irvine Machine Learning Repository	Currently maintaining 557 data sets as a service to the machine learning community
Biomedical & Life Sciences Datasets	Library resources in Biology by Brown University
Awesome Public Datasets	A list of a topic-centric public data sources in high quality, collected from blogs, answers, and user responses
Kaggle Datasets	Containing over 19,000 public datasets and 200,000 public notebooks to conquer any analysis in no time
The home of the U.S. Government's open data	Data, tools, and resources to conduct research, develop web, mobile applications and design data visualizations
Bureau of Labor Statistics	BLS Public Data gives the public access to raw economic data from all BLS programs
Amazon Web Services (AWS) datasets	Help people discover and share datasets that are available via AWS resources
BigQuery Public Datasets	Public datasets are available for you to analyze using either legacy SQL or standard SQL queries



YouTube-8M Segments Dataset

A large-scale labeled video dataset that consists of millions of YouTube video IDs, with high-quality machine-generated annotations from a diverse vocabulary of 3,800+ visual entities

Harvard Dataverse

A repository for research data by Harvard University

Large Health Data Sets

A freely accessible repository for healthcare data

World Bank Open Data

Free and open access to global development data

Popular Open Source Datasets

Popular open source and public data sets, data visualization, data analytics and data lakes