Finding Data Sources

Recommended Data Collections

Whether you are looking for research data, or looking to practice your data science skills, below are some data collections we recommend.

Source	Description	Access and Notes
UC Irvine Machine Learning Repository	The UCI Machine Learning Repository is a collection of databases, domain theories, and data generators that are used by the machine learning community for the empirical analysis of machine learning algorithms.	Managed and hosted by UC Irvine. Free and open access.
Kaggle.com	Kaggle is an online community that hosts competitions for machine learning. Kaggle also holds a collection of datasets, open for public use, although it requires users to create an account. Kaggle was acquired by Google in 2017.	Requires user email address to create an account.
Tidy Tuesday	A weekly data project aimed at the R programming language ecosystem. As this project was borne out of the `R4DS Online Learning Community` and the `R for Data Science` textbook, an emphasis was placed on understanding how to summarize and arrange data to make meaningful charts with `ggplot2`, `tidyr`, `dplyr`, and other tools in the **`tidyverse`** ecosystem. Datasets are posted each week and available on the R for Data Science GitHub.	Free and open access. UCLA Library Data Science Center periodically hosts Tidy Tuesday welcoming all to practice working with data together.
UCLA Dataverse	Collection of research data and findings from the UCLA research community.	Free and open access.
GitHub Awesome Public Datasets	Curated lists of public datasets on GitHub.	Free and open access.

Additional Information

Data Science Center

Submit a Request

See also

- Dataverse
- Data Science Center
- Geospatial services

- 3D printing
- 3D scanning
- Accounts & Access

This service is available to

All	
Library Staff	✓
UCLA	✓
Campus-Affiliated	
Other	

If this service is not available to you, and you have questions, please send an email to techhelp@library.ucla.edu for more information.

Please visit the Data Science Center portal to raise a request.

- API (Application Programming Interface)
- Apple/Mac
- ArticlesPlus
- Automated security patching
- Bar codes for books
- BruinPrint/wepa kiosks
- BSLW (Backstage Library Works) bibliographic load
- Carpentries Program
- Catalog & Library Systems
- CCLE (Common Collaboration and Learning Environment)
- CLICC classrooms
- CLICC fines
- clicc.library.ucla.edu - CLICC Virtual Desktops
- Cloud printing
- Communication & Collaboration
- Computers
- Content Management System (CMS)
- Copiers
- Course Management System (CMS)
- Course reserves
- Data & Coding Support
- Data Science Center
- Data Services
- Database Administration
- Departmental drives
- Desktop applications
- Dissertation embargoes
- DocuSign
- Drupal
- E-waste service
- Email
- Equipment loans
- Equipment provisioning
- eScholarship
- ETD (Electronic Thesis Dissertation) load
- File document restoration /backup
- File storage
- Finding Data Sources
- Geospatial services (GIS)
- Google Analytics
- Hardware

- Instructional Technology & Reserves
- Internet & Network
- Internet connection
- Intota
- JasperReports
- Laptop lending
- Laser etching & laser cutting
- Library domain
- Load Course Reserves in to CCLE
- Lynda.com
- Machine Learning /Artificial Intelligence (ML /AI)
- Macro Express
- MARCIVE GovDocs load
- Moving and relocation
- OCLC daily load
- Oversize printing
- Printers
- Printing
- Printing & Scanning
- Project Management Services - DIIT
- PromptCat load
- Publication Management System
- Remote access
- Remote desktop
- Search Engine Optimization (SEO)
- Security
- SIA (Summary of Instructional Activity) database
- Software
- Software configuration
- Storage & Cloud Services
- Study rooms
- Training & Consultation
- UCLA Dataverse
- UCLA Library Catalog
- Usability
- Voyager
- Voyager account
- Voyager user privileges
- Web accessibility
- Web design
- Web service implementation
- WiFi

- Windows Server Update Services Server (WSUS)
- Wireframing
- Zoom