

Generic Public Datasets

- [Kaggle](#)
- [Google Dataset Search](#) - Dataset search engine
- [BigQuery Datasets](#) - Datasets publicly available on BigQuery and managed by Google
- [BigQuery Datasets](#) - Datasets publicly available on BigQuery
- [Google Trends](#)
- [UC Irvine Datasets](#)
- [538 datasets](#)
- [Buzzfeed open datasets](#)
- [BU Library resources](#)

Recommender system datasets (movies, food, products, dating, etc.):

<https://gist.github.com/entaro/adun1653794>

Data.gov

<https://www.data.gov/>

World Bank Open Data

<https://data.worldbank.org/>

Google Cloud Public Datasets

<https://cloud.google.com/public-datasets>

Yelp Dataset Challenge

<https://www.yelp.com/dataset>

Amazon Web Services Public Datasets

<https://registry.opendata.aws/>

National Voter File

<https://www.bu.edu/cds-faculty/2021/12/16/bu-announces-new-data-source-national-voter-file/>

Historical Sales and Active Inventory

<https://www.kaggle.com/flenderson/sales-analysis>

2015 Flight Delays and Cancellations

<https://www.kaggle.com/usdot/flight-delays>

Airlines Delay

<https://www.kaggle.com/giovamata/airlinedelaycauses>

Uber Pickups in New York City

<https://www.kaggle.com/fivethirtyeight/uber-pickups-in-new-york-city>

Telco Customer Churn

<https://www.kaggle.com/blastchar/telco-customer-churn>

Craft Beers Dataset

<https://www.kaggle.com/nickhould/craft-cans>

Retail Data Analytics

<https://www.kaggle.com/manjeetsingh/retaildataset>

E-Commerce Data

<https://www.kaggle.com/carrie1/ecommerce-data>

Brazilian E-Commerce Public Dataset by Olist

<https://www.kaggle.com/olistbr/brazilian-ecommerce>

Global Garment Data

<https://old.datahub.io/dataset/global-garment-supply-chain-data>