

Kaggle | <https://www.kaggle.com/datasets>

### **1. Trending YouTube Video Statistics**

Possible uses for this dataset could include:

- Sentiment analysis in a variety of forms.
- Categorizing YouTube videos based on their comments and statistics.
- Training ML algorithms like RNNs to generate their own YouTube comments.
- Analyzing what factors affect how popular a YouTube video will be.
- Statistical analysis over time.

<https://www.kaggle.com/datasnaek/youtube-new>

### **2. Open Food Facts**

Open Food Facts is a free, open, collaborative database of food products from around the world, with ingredients, allergens, nutrition facts and all the tidbits of information we can find on product labels.

<https://www.kaggle.com/openfoodfacts/world-food-facts>

### **3. World Happiness Report**

Happiness scored according to economic production, social support, etc.

<https://www.kaggle.com/unsdsn/world-happiness>

### **4. Where it Pays to Attend College**

Salaries by college, region, and academic major

<https://www.kaggle.com/wsj/college-salaries>

### **5. European Soccer Database**

25k+ matches, players & teams attributes for European Professional Football

<https://www.kaggle.com/hugomathien/soccer>

## **6.Global Terrorism Database**

More than 170,000 terrorist attacks worldwide, 1970-2016

<https://www.kaggle.com/START-UMD/gtd>

[NYC OpenData | https://opendata.cityofnewyork.us/](https://opendata.cityofnewyork.us/)

### **1. DOB Job Application Filings**

A list of job applications filed for a particular day and associated data. Prior weekly and monthly reports are archived at DOB and are not available on NYC Open Data.

<https://data.cityofnewyork.us/Housing-Development/DOB-Job-Application-Filings/ic3t-wcy2>

### **2. DOHMH New York City Restaurant Inspection Results**

This dataset provides restaurant inspections, violations, grades and adjudication information.

<https://data.cityofnewyork.us/Health/DOHMH-New-York-City-Restaurant-Inspection-Results/xx67-kt59>

### **3. MTA Data**

Information pertaining to MTA (Metropolitan Transportation Authority of the State of New York) subways, buses, commuter rail, bridges, and tunnels

<https://data.cityofnewyork.us/Transportation/MTA-Data/mmu8-8w8b>

### **4. NYC Jobs**

This dataset contains current job postings available on the City of New York's official jobs site (<http://www.nyc.gov/html/careers/html/search/search.shtml>). Internal postings available to city employees and external postings available to the general public are included.

### **5. New York City Leading Causes of Death**

The leading causes of death by sex and ethnicity in New York City in since 2007.

<https://data.cityofnewyork.us/Health/New-York-City-Leading-Causes-of-Death/jb7j-dtam>

## **6. NYC Business Acceleration Businesses Served and Jobs Created**

The list tracks the number of businesses that NYC Business Acceleration has assisted in opening and how many jobs were created by those businesses. This data is up to date as of the date reflected in the "About" tab of this dataset.

<https://data.cityofnewyork.us/Business/NYC-Business-Acceleration-Businesses-Served-and-Jo/9b9u-8989>

DATA.GOV | <https://www.data.gov/>

### **1. Expenditures on Children by Families**

This dataset provides expenditures on Children by Families provides estimates of the cost of raising children from birth through age 17 for major budgetary components.

<https://catalog.data.gov/dataset/expenditures-on-children-by-families>

### **2. Housing Affordability Data System (HADS)**

These datasets give the community of housing analysts the opportunity to use a consistent set of affordability measures.

<https://catalog.data.gov/dataset/housing-affordability-data-system-hads>

### **3. U.S. International Trade in Goods and Services**

Monthly report that provides national trade data including imports, exports, and balance of payments for goods and services.

<https://catalog.data.gov/dataset/us-international-trade-in-goods-and-services>

### **4. Food Environment Atlas**

Food environment factors--such as store/restaurant proximity, food prices, food and nutrition assistance programs, and community characteristics--interact to influence food choices and diet quality.

<https://catalog.data.gov/dataset/food-environment-atlas-f4a22>

### **5. Economic Census**

Department of Commerce — Provides a detailed portrait of business activities in industries and communities once every five years, from the national to the local level.

<https://catalog.data.gov/dataset/economic-census>

The relationship between weather and the level of happiness

- <https://www.kaggle.com/unsdsn/world-happiness>
- weather

The relationship between the favorite type of restaurant and the community

- <https://catalog.data.gov/dataset/food-environment-atlas-f4a22>
- region-community

The potential impact of war trade

- <https://catalog.data.gov/dataset/us-international-trade-in-goods-and-services>
- the estimated ratio between GPA and import, output