Data Visualization: 7th lesson – Final Project

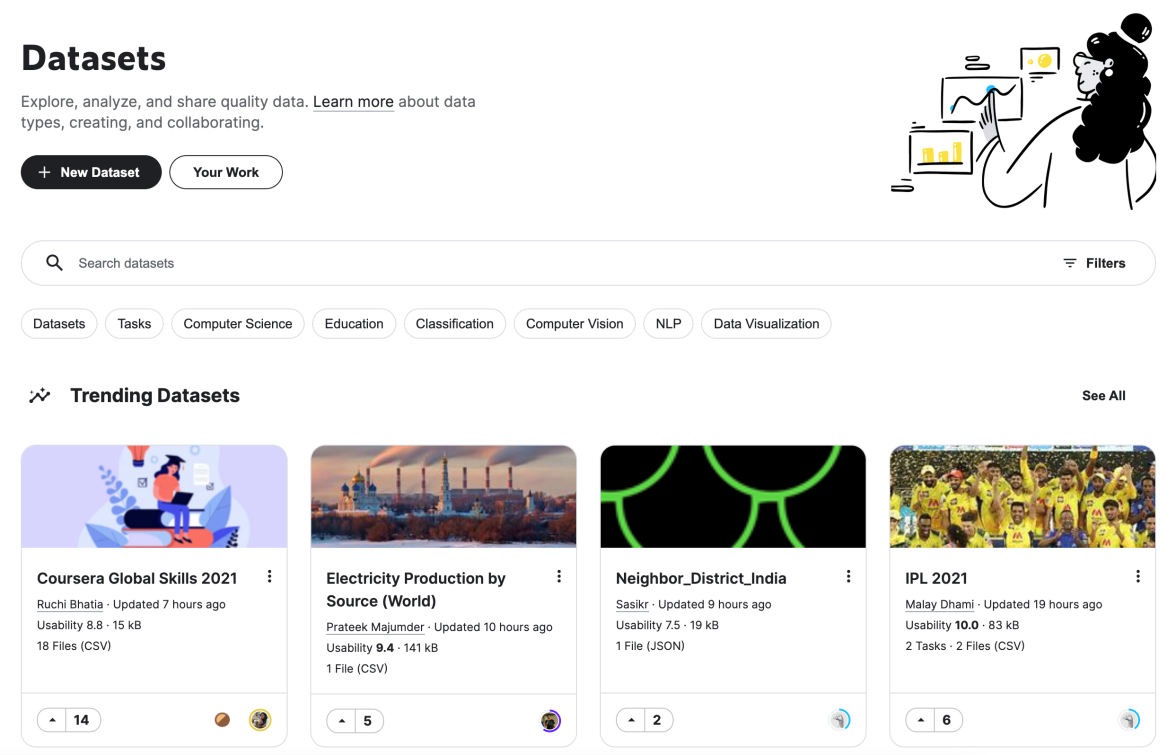
So far, you have worked with datasets that we have provided for you. In this tutorial, you'll learn how to use your own datasets. Then, in the following exercise, you'll design and create your own data visualizations.

You'll learn all about Kaggle Datasets, a tool that you can use to store your own datasets and quickly access tens of thousands of publicly available data sources.

***Kaggle Datasets***

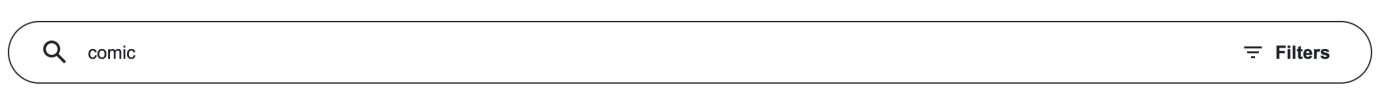
You can access Kaggle Datasets by visiting this link: <https://www.kaggle.com/datasets>

The link will bring you to a webpage with a collection of datasets that you can use in your own projects.

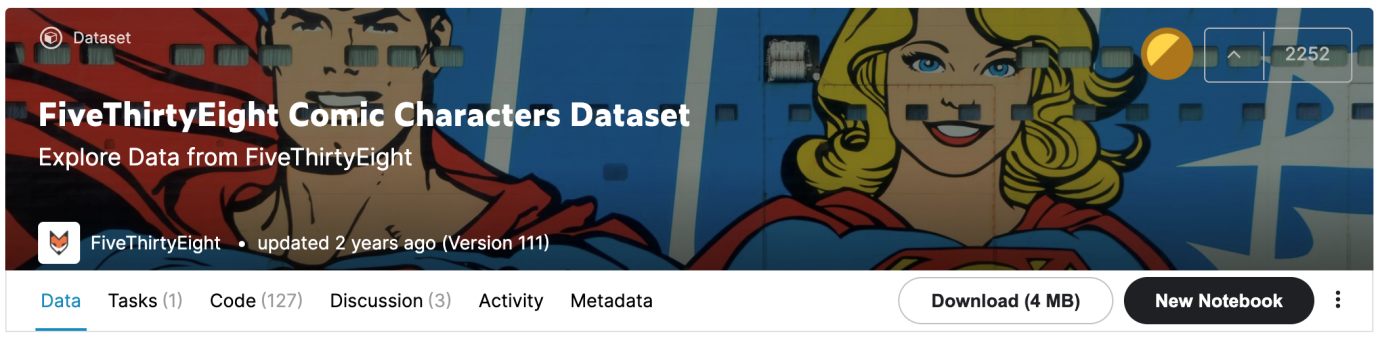


Note that the datasets that you see on the page will likely look different from what's shown in the screenshot above, since many new datasets are uploaded every day! There are many different file types on Kaggle Datasets, including CSV files, but also more exotic file types such as JSON, SQLite, and BigQuery. We'll be careful to select a dataset with at least one CSV file, since that is the file type we have been working with in this course.

To search for a specific dataset, use the search bar at the top of the screen. Say, for instance, you'd like to work with a dataset about comic book characters. Begin by typing "comic" in the search window.



Then, find the *FiveThirtyEight Comic Characters Dataset*. Note that the dataset contains 3 files, including a CSV file that we can use. Then, click on the dataset to select it. This will bring you to a webpage that describes the dataset.



Scroll down to see the list of files in the dataset under Data Explorer, on the left of the window. The dataset contains three files:

* README.md
* dc-wikia-data.csv
* marvel-wikia-data.csv

The first file is selected as default. Click on one of the CSV files instead to see a quick preview of the file. Take the time now to explore the other tabs on the page; for instance, check out the Discussion tab to see what others have to say about the dataset.

***Use your own dataset***

In the following exercise, you'll work with any CSV dataset of your choosing. As you learned above, Kaggle Datasets contains a large collection of datasets that you can use. If you'd prefer to work with your own dataset, you'll need to first upload it to Kaggle Datasets.

To do that, return to kaggle.com/datasets, and click on the "New Dataset" button in the top left corner. If you're not familiar with CSV file types, note that the Kaggle Datasets platform will automatically convert any tabular data that you have to a CSV file. So, feel free to upload something like a Google spreadsheet or an Excel worksheet, and it will be transformed to a CSV file for you!