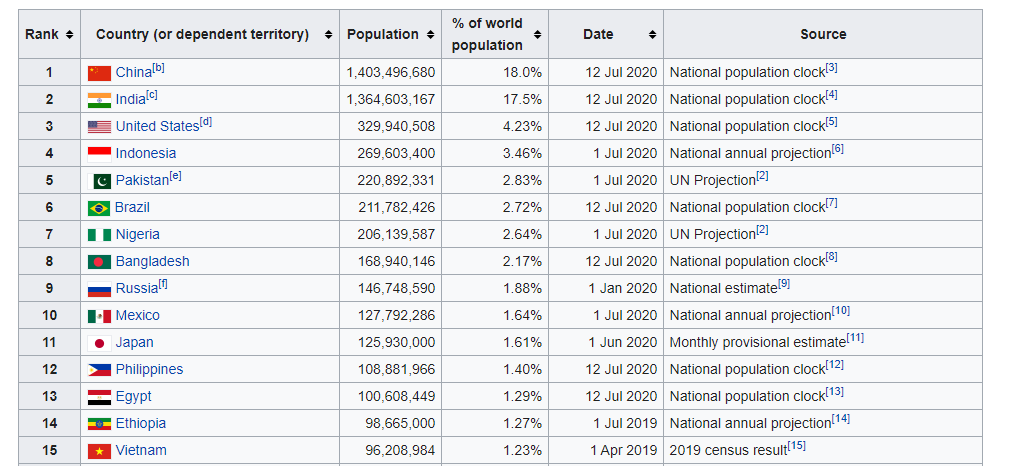
**Milestone 1**

*Identify Datasets*

The first milestone of this project will be to select the data you want to work with. You will need to select 3 different data sources that have different file types of information – and the data will need to have a relationship between them. If one doesn’t exist, you will have to create one. It is likely you will need to manipulate the data to create a relationship. Finding the data, you want to work with for this project, will likely be the hardest part of the project. You must have one of each of the following types of datasets – and you need a minimum of 1000 rows across all datasets. You need a total of 30 columns across the 3 datasets you select.

* CSV/Excel/PDF or another flat file source.
* Website you want to pull data from--you will want to identify a website that has data stored in a table, similar to the screenshot below.



* API you will pull data from.

Some places you can find datasets are listed below:

* [Tableau Community](https://community.tableau.com/docs/DOC-10635)
* [Kaggle Datasets](https://www.kaggle.com/datasets)
* [Data.Gov](http://www.data.gov/)
* [Science.Gov](http://www.science.gov/)
* [Data.Gov.UK](http://data.gov.uk/)
* [NORC](http://gss.norc.org/)
* [European Social Survey](http://www.europeansocialsurvey.org/)
* [API List](https://apilist.fun/)
* [PrommableWeb](https://www.programmableweb.com/apis/directory)
* [Public APIs](https://github.com/toddmotto/public-apis/blob/master/README.md)
* [OpenWeatherMap](https://openweathermap.org/api)

Wikipedia is a good source to find data that is in a table - and the structure of the HTML is usually very similar.

There are no restrictions on what dataset you use, other than you cannot use the specific datasets used in the book(s).

For the first milestone, you need to submit the following:

* Project Subject Area: Describe your project in 1-2 sentences
* Data Sources:
  + Flat File:
    - Description
    - Link or Flat File uploaded
  + API:
    - Description
    - Link
  + Website:
    - Description
    - Link
* Relationships
  + Describe how the data from each source is connected (see example below).
  + If there isn't an obvious relationship, explain how you will make one
* 250 Words describing how you plan to tackle the project, what the data means, ethical implications of your project scenario/topic, and what challenges you might face.

Submit via a PDF to the assignment link.

**Example of Relationships:**

In case you are confused what is meant by a relationship between the data sources here is an example (this is a very simple example and I would expect your datasets to have more variables)

CSV File: Contains a list of stores by store ID and other metadata about the stores

Website: Contains a list of store locations, by location ID and store ID and the various departments each store has by department ID.

API: Contains the transactions at each store – contains a transaction ID and store ID.

All 3 of these data sources are related by Store ID. The CSV file has a 1 to many relationship with the Website by StoreID and has a one to many relationship with the API data by StoreID as well.