There should be a file called read.py already open. You can run this from the command line by being in the same folder, and typing python read.py. Of course, there's nothing in the file right now. You might recall from the last mission that you can put this into a file to run it from the command line:



if \_\_name\_\_ == "\_\_main\_\_":

   print("Welcome to a Python script")

This will print Welcome to a Python script on the command line if you put it into a file and run it.

We can also add functions into a file by writing them like normal:



def load\_data():

   pass

​

if \_\_name\_\_ == "\_\_main\_\_":

  # This will call load\_data if you run the script from the command line.

  load\_data()

Function definitions should come before the if \_\_name\_\_ == "\_\_main\_\_" line. These functions can be imported from other files.

We'll be adding some code to the read.py file that will help us load in the dataset and do some initial processing. We'll then be able to import the code to read in the dataset from other scripts we develop.

Instructions

* In the read.py file, read the hn\_stories.csv file into a Pandas Dataframe.
* There is no header row in the data, so the columns don't have names. See [this stackoverflow thread](https://stackoverflow.com/questions/34091877/how-to-add-header-row-to-a-pandas-dataframe)for how to add column names. Add the column names from the last screen (submission\_time, upvotes, url, and headline) to the Dataframe.
* Create a function called load\_data that takes no inputs, but contains the code to read in and process the dataset. load\_data should return a Pandas Dataframe with the column names set correctly.

As you work on these steps, you should be running your script on the command line every so often and verifying that things are working. You can run read.py from the command line by calling python read.py. The first verification is to make sure that you don't see any errors. The second one is to call print at key points in your code, and make sure that the output looks like what you expect. You might want to do this after each step above. This is a good general rule of thumb to follow when writing new code.