* Step0: You need to identify the URL that you want to scrape information from
  + Conceptually, for a single page this is easy.
  + If you need to scrape information from multiple webpages, you will need to understand how the URL changes so you can programmatically account for that.
* Step1 : get the entire contents of the HTML page into your software of choice (we will be using R) – from being nice perspective, you should do this once for every webpage
* Step2:
  + For every “thing” you are interested in, we will need to figure out the HTML elements of interest.
  + For example, for the imdb page, you may be interested in **titles, scores, and summary of the movie/show** 🡪 for each of these three “things”, you will need to find the HTML elements that you will be using for scraping.
* Step3: Would be to clean out the html tags and include only the human readable content in our output.
* Step4: Saving everything that we got in dataframe/tibble so we can easily export that to a flat file (CSV).