In class you saw how you can use R to scrape data from web pages and also analyze that data. We used IMDB, the Internet Movie DataBase as a data source and calculate average ratings and average number of votes for the top/bottom-rated movies on the site and also created a wordcloud based on the short summaries of a select number of movies.

But what if you are interested specifically in Sci-Fi movies (and, to be honest, who isn't)? And, as an aspiring filmmaker, you would like to find out what the average runtime of the top-rated Sci-Fi movies is so that you can make sure that your own work is going to hit exactly that mark?

This is where this exercise comes into play. Based on what you learned about web scraping with R, you will create a short script (muuuuch shorter than the one we used in class) that will calculate the average runtime of all movies listed on the first page of IMDB's list of "Highest Rated Sci-Fi Feature Films With At Least 25000 Votes." You can find the page at the following address (make sure to get the comma at the very end):

http://www.imdb.com/search/title?genres=sci_fi&sort=user_rating,desc&title_type=feature&num_votes=25000,

To achieve this goal, think about the following:

- Which libraries do you need to achieve your goal?
- Where on the page do you find the information you need, i.e., which element with which class value is used to contain that information?
- If you extract that piece of content, does the string contain elements that you don't need or that can even cause problems later on? How do you get rid of them?
- What type of data is it, i.e., is it a character string, a number, etc.? What type do you need to do your calculation?
- How do you calculate your average runtime?
- **BONUS:** Your average runtime is likely going to have a couple of decimal points. How do you get rid of them? (Hint: The programmer's best friends are Google and documentation).

When you run your script, it should print out a line that says something like "Average Runtime: 25 minutes".

One possible solution with explanations is shown in the file web_scraping_exercise.R.