

Setting up or “Hello World”

The first problem set is aimed at making sure you are comfortably set up, installing R on your machine as well as setting up Markdown. Homework assignments will be submitted in form of Markdown exported PDF document.

- 1) Download R-Studio and install it on your laptop.
<https://www.rstudio.com/products/rstudio/download/>
- 2) Make yourself familiar with R-studio and create a new markdown document with a headline, author information and date of creation.
- 3) Install the **stringr** r-package as a tool for string manipulation.
(`install.packages("stringr")`)
- 4) Using the **readLines()** command, grab the content of <https://www.nytimes.com/> and store as object. Print out first rows using the `head()` function.
- 5) Using **grep** to subset the NYT object to only include the parts visible in the browser window (look up the meaning of the body-HTML Tag). Save the result in a new object and print the first 10 rows.
- 6) Using **gsub** remove all HTML and print the first lines of the cleaned item.
- 7) Extract all the **h2 headlines**, save them as a list and print them.

TIP: You can access the source code of a website using the developer mode of your navigator.

Send your PDF version of the Markdown document to

thiemo_837e@sendtodropbox.com

Please use the following naming convention: HW1_Surname_FirstName.pdf

Make sure to send the email from your @uchicago or @chicagobooth email address as only these domains are white-listed.

Deadline: by Wed 12th April, 2017.