

- **NO** late submission will be accepted, except under special circumstances.
 - Homework must be done individually and not in groups. Discussion of problems with others is permitted (and encouraged!), but you must write your own work in your own words.
 - Submit your answers (via Canvas) as a single RMarkdown file that can be run on anyone's machine (i.e., that doesn't refer to your local files or directories). Your file name should have the following format: `lastname-NetID-week07.Rmd`. Make sure that your Rmarkdown file has yourself as author and has `output:html_document`.
 - Be sure to include detailed explanatory text and remarks of what you are doing—don't just show a lot of R code and computer generated output. Use commands from the `tidyverse` and pipes whenever you can.
1. Read Chapter 25, “Many Models”, in *R for Data Science*.
 2. Look at the API documentation at <https://dev.socrata.com/foundry/healthdata.nj.gov/9hse-wixk>.
 - (a) Write a function that will use the API and then generate a plot of the rate of heart disease over time by year. Set the default to all races, but include an option to specify the race. To get `ggplot2` to generate a plot from within a function, wrap the final object in the `print()` function. Be sure to include in your RMarkdown a plot generated by your function, using the default of all races and another using one specific race.
 3. The *New York Times* has a nice set of APIs, described at <https://developer.nytimes.com/apis>.
 - (a) Get yourself an API key.
 - (b) Make 2 barplots, one with the most common non-stop-words in the titles of the Most Popular articles by views for the past week (<https://developer.nytimes.com/docs/most-popular-product/1/overview>) and another of the most common non-stop-words in the titles of the “world” Top Stories articles (<https://developer.nytimes.com/docs/top-stories-product/1/overview>).
 - (c) IMPORTANT: Do **not** include your API key in your RMarkdown file. Instead, create a file called `api-keys.R` to store the key as a string called `api.key.nytimes`, and include `source("../api-keys.R")` in the preamble to load the file.
 4. Taking advantage of the `rvest` package, turn the table at <http://www.nature.com/articles/ng.3097/tables/3> into an R data frame.
 - (a) Be sure to delete the rows “Genes with previous literature support (GRAIL)” and “New genes without previous evidence” (and don't do it by using the row number).
 - (b) Be sure to convert the *p*-value column to numbers.

- (c) The last column is shown in 3 rows in the journal, but most likely as one string in your table. Use regular expressions to insert semicolons (i.e., “;”) between each of the original lines. For example,
- “PI3K cascade (REACTOME, $P = 6.2 \times 10^{-13}$); Chronic myeloid leukemia (KEGG, $P = 1.6 \times 10^{-12}$); Response to fibroblast growth factor stimulus (GO, $P = 5.4 \times 10^{-11}$)”
- (d) Show the table in your RMarkdown file by using the function `knitr::kable()`.