MSDA 607 - Week 9 Assignment

Your Week 9 assignment is due end of day on Sunday March 29th. *This is a warm up exercise to help you to get more familiar with the HTML, XML, and JSON file formats, and using packages to read these data formats for downstream use in R data frames. In the next two class weeks, we'll be loading these file formats from the web, using web scraping and web APIs.*

⌂

Pick three of your favorite books on one of your favorite subjects. At least one of the books should have more than one author. For each book, include the title, authors, and two or three other attributes that you find interesting.

Take the information that you've selected about these three books, and separately create three files which store the book's information in HTML (using an html table), XML, and JSON formats (e.g. "books.html", "books.xml", and "books.json"). To help you better understand the different file structures, I'd prefer that you create each of these files "by hand" unless you're already very comfortable with the file formats.

Write R code, using your packages of choice, to load the information from each of the three sources into separate R data frames. Are the three data frames identical?

Your deliverable is the three source files and the R code. If you can, package your assignment solution up into an .Rmd file and publish to rpubs.com. [This will also require finding a way to make your three text files accessible from the web].