**Homework #3 (Due Sep 17th 11:59 PM)**

IST 3420 - Fall 2017, Chen

**Name**: \_\_\_\_\_Adam Forestier\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Programming Task (20 points)**

**Instruction:** Follow below major steps to complete the programming task. Provide answers in boxes. Upload this document with your answers and your R script file to “Homework 3” on Canvas.

1. Read an HTML table from <https://nrf.com/2015/top100-table>. The table summarize top 100 retailers in 2015.
2. Use a data frame named “retailer\_df” to store the data collected from the above HTML table.
3. Show the attributes and structure of the data frame. List all original column names in the following box.

|  |
| --- |
| **Your Answer** (**2 points**):  1. Top 100 Rank 2. Company 3. Headquarters/U.S. Headquarters 4. 2014 Retail Sales (000) 5. USA Sales Growth (’14 v ’13) 6. Worldwide Retail Sales (000) 7. USA % of Worldwide Sales (000) 8. 2014 Stores 9. Growth(’14 v ’13) |

1. Rename the columns as:
   1. Rank
   2. Company
   3. Headquarter
   4. RetailSales2014
   5. USASalesGrowth
   6. WorldwideRetailSales
   7. USAPercentageOfWorldwideSales
   8. Stores2014
   9. Growth
2. Show summary statistics of the dataset. Do you get the min, median, mean, and max values of the RetailSales2014 variable? Why? Put your answers in the following box.

|  |
| --- |
| **Your Answers** (**2 points**):  No, its class is currently character. It needs to be changed to an integer or numeric to be able to show statistics. |

1. Save the data frame as a CSV file “top100retailers2015.csv”.
2. Paste your R script code in the following box.

|  |
| --- |
| **Your R Code** (**14 points**):  # Name: Adam Forestier  # Class: IST 3420  # Assignment: Homework 2  # Date: September 17, 2017  # Package that reads HTML Table  install.packages("rvest")  library(rvest)  # Sets url to variable url  url <- "https://nrf.com/2015/top100-table"  # Reads HTML File and makes a table with data  read\_retail <- read\_html(url)  retail\_table <- html\_table(read\_retail, fill = TRUE)[[1]]  # Converts Table to Data Frame  retailer\_df <- as.data.frame(retail\_table)  retailer\_df  # Shows attributes of data frame  attributes(retailer\_df)  # Shows data structure of data frame  str(retailer\_df)  # Renames data frame columns  names(retailer\_df) <- c("Rank", "Company", "Headquarter", "RetailSales2014", "USASalesGrowth",  "WorldwideRetailSales", "USAPercentageofWorldwideSales", "Stores2014", "Growth")  # Shows summary of data frame  summary(retailer\_df)  # Saves data frame as a CSV file  write.csv(retailer\_df, file = "top100retailers2015.csv") |

1. Upload this documents with your answers to “Homework 3” on Canvas.
2. Upload your R script file to “Homework 3” on Canvas. (**2 points**)