Problem Set 4

savannahjsimpson

February 20 2024

1 Web Scraping

- Prior to this course, I was not familiar with the parameters of web scraping. Since learning the process and realizing how useful this could be, I have had thoughts about how I could use this tool. I am not yet completely sure what my project will follow this semester, but I am considering using FRED data. I found some interesting data on the Florida home pricing index that I would like to use. Additionally, I am currently doing research on suggested tipping. Scraped Twitter data including specific keywords could also be beneficial for additional public analysis.

2 Problem Set Questions

- (Question 5d) Object mydf is a table data frame. Object mydfdate is a character in R.

- (Question 6:7) Object df1 is a table data frame. Object df is a Spark table data frame but also a SQL file, that operates with a lazy tibble.

- (Question 6:8) The objects' column titles are the same terms, but one is separated by periods while the later is separated by underscores. While they have similar formation, the different types of files will carry different formats.