

Problem Set 4

Trevor Jones

February 21, 2023

1 Questions

1. tell me about some data sources that you would be interested in scraping from
 - (a) I would like to scrape data from some classical texts and do some work with them. One project I've been interested in doing since I took Stats for Decision making with Dr. Wang is investigating classical texts with disputed authors to see if there is any statistical analysis that can give light to certain disputes. An example of this is Homer's authorship of the Iliad and the Odyssey (these are in Greek, but along the lines of what interests me).
2. In another part of your .tex file, answer the questions raised in the various parts of the previous question
 - (a) output on class of df1: [1] "tbl.df" "tbl" "data.frame"
 - (b) output on class of df2: "tbl.spark" "tbl.sql" "tbl.lazy" "tbl"
 - (c) the column names are the same, but they one uses an underscore between words while the other uses a dot.