STAT 847: Reading Assignment 1

DUE: Friday January 19, 2024 by 11:59pm Eastern

NOTES

Your assignment must be submitted by the due date listed at the top of this document, and it must be submitted electronically in .pdf format via Crowdmark.

Organization and comprehensibility is part of a full solution. Consequently, points will be deducted for solutions that are not organized and incomprehensible. Furthermore, if you submit your assignment to Crowdmark, but you do so incorrectly in any way (e.g., you upload your Question 2 solution in the Question 1 box), you will receive a 5% deduction (i.e., 5% of the assignment's point total will be deducted from your point total).

Reading: Hands-On Exploratory Data Analysis with R [8 marks]

Open the UWaterloo Library website, lib.uwaterloo.ca , and use your WatIAM account to search for an open the book Hands-On Exploratory Data Analysis with R. By Radhika Datar, Harish Garg. The following questions can be answered by reading the "Univariate and Control Datasets" chapter.

Please put your answers to questions 1-4 on a separate page from your answers to questions 5-8, this can be done in Word with Crtl + Enter, or in Markdown with `\newpage`.

Each question is work one mark.

(Unless "in your own words" is specified, you can directly quote the book.)

Winter 2024 Reading Assignment

Questions

Q1. What's the name of the test for outliers used in this chapter?

Q2. What does the variable 'pdays' represent?

Q3. How many rows are there in the bank marketing data?

Q4. What does each row represent in the bank marketing data?

Q5. (Challenge) What is two sample t-test actually comparing in the "the t-test in R" page?

Q6. What makes a model parsimonious?

Q7. Almost every named distribution (e.g., the normal, the uniform) has a function that calculates its cumulative distribution function. What is the letter that all such functions start with?

Q8. According to the Shapiro-Wilk test, are bank balances normally distributed?