MSIS545L 232-711 Prof. Gissella Bejarano
Spring 2024 - Assignment 4 (100 pts)
(based on The Statistical Sleuth)
Due date April, Sunday 14th – 11:59pm

No need to rewrite the exercise' statements. Work in a text processor, distinguish the font of your analysis from the R script, and convert your document to PDF. Include proper plots and scripts to support your analysis.

WRITTEN PART

Show steps in your calculations and procedures.

1. **Comparison of Five Teaching Methods – (40 pts)**

   An article reported the results of a planned experiment contrasting five different teaching methods. Forty-five students were randomly allocated, nine to each method. After completing the experimental course, a one-hour examination was administered. Table 1 summarizes the score on a 10-minute retention test that was given 6 weeks later. (Data from S.W. Tsai and N. F. Pohl, "Computer-Assisted Instruction Augmented with Planned Teacher/Student Contacts" – Journal of Experimental Education 49(2))

Group	Logo	Teaching Method	n	Average	SD
1	L+D	Lecture and Discussion	9	30.2	3.82
2	R	Programmed text	9	28.8	5.26
3	R+L	Programmed text with lectures	9	26.2	4.66
4	C	Computer instruction	9	31.1	4.91
5	C+L	Computer instruction with lectures	9	30.2	3.53

Table 1. Test scores for the experimental CAD instruction course

   A) Compute the pooled estimate of the standard deviation from these summary statistics.
   B) Determine a set of coefficients that will contrast the methods using programmed text as part of the method with those that do not use programmed text. Make sure each super group is compared and added up to 1 or -1.
   C) Evaluate the comparison by calculating the g statistic using (b) and compute a 95% confidence interval.

CODING PART

2. Education and Future Income (60 pts)

The data file ex0525 contains the annual incomes in 2005 of a random sample of 2,584 Americans who were selected for the National Longitudinal Survey of Youth in 1979 and who had playing jobs in 2005. The data set also includes a code for the number of years of education that each individual had completed by 2006: <12, 12, 13-15, 16, and >16.

A) How strong is the evidence that at least one of the five population distributions (corresponding to different years of education) is different from the others?
B) Select a multiple comparison procedure and report which groups are different.
C) Visually analyze the data and select a proper transformation. Perform the same task in b) and report if the evidence in a) changes.
D) By how many dollars or by what percent does the mean or median for each of the last four categories exceed that of the next lowest category?
E) Assuming that every group can be related to the additional explanatory variable such as the average number of years studied in their groups: 10, 12, 14, 16, 20. Is there any evidence for a linear trend in income as a function of the number of years studied?