BACS HW – Week 7

The data in this assignment is from an actual study but the scenario is fictional:
A health researcher is investigating how health information spreads through word-of-mouth across different media (video, pictures, text, etc.). She has prepared some informative content about avoiding stomach aches. She is curious which media format to use, or avoid, in order to encourage people to share such health related information. She wants to conduct her tests at 95% confidence.

So she prepares the similar information content in four alternative media formats:
(1) *video [animation + audio]: A fully animated video with audio narration*
(2) *video [pictures + audio]: Video of sequence of still pictures (not-animated) with audio narration*
(3) *webpage [pictures + text]: Static web page with still pictures (no video) and accompanying text narration (no audio)*
(4) *webpage [text only]: Static web page of text narration (no audio) but no pictures*

Our researcher runs an experiment where each of these four alternative media is shown to a different panel of randomly assigned people. Afterwards, viewers were surveyed about their thoughts, including a question (labeled *INTEND.0*) about their intention to share what they had seen with others:
[INTEND.0]: I intend to share the information I saw with others.
(answered on 7 point scale: 1=strongly disagree; 4=neutral; 7=strongly agree)

You may find the researcher's data in a ZIP file containing four CSV files named: *pls-media[1-4].csv*
(note: the number in the filename corresponds to the type of media listed above)

Question 1) Let's **develop some intuition** about the data and results:

a. What are the *means* of viewers' intentions to share (INTEND.0) on each of the four media types?
b. Visualize the *distribution and mean* of intention to share, across all four media.
   (Your choice of data visualization; Try to put them all on the same plot and make it look sensible)
c. From the visualization alone, do you feel that media type makes a difference on intention to share?

Question 2) Let's try **traditional one-way ANOVA**:

a. State the null and alternative hypotheses when comparing INTEND.0 across four groups in ANOVA
b. Let's compute the F-statistic ourselves:
   i. Show the code and results of computing MSTR, MSE, and F
   ii. Compute the p-value of F, from the null F-distribution; is the F-value significant?
   If so, state your conclusion for the hypotheses.
c. Conduct the same one-way ANOVA using the aov() function in R – confirm that you got similar results.
d. Regardless of your conclusions, conduct a post-hoc Tukey test (feel free to use the TukeyHSD() function in R) to see if any *pairs of media have significantly different means* – what do you find?
e. Do you feel the classic requirements of one-way ANOVA were met?
   (Feel free to use any combination of methods we saw in class or any analysis we haven't covered)

See next page for question 3

Question 3) Let's use the **non-parametric Kruskal Wallis** test:

a. State the null and alternative hypotheses (in terms of distribution or difference of mean ranks)
b. Let's compute (an approximate) Kruskal Wallis H ourselves:
    i. Show the code and results of computing H
    ii. Compute the p-value of H, from the null chi-square distribution; is the H value significant? If so, state your conclusion of the hypotheses.
c. Conduct the same test using the kruskal.wallis() function in R – confirm that you got similar results.
d. Regardless of your conclusions, conduct a post-hoc Dunn test (feel free to use the dunnTest() function from the FSA package) to see if any *pairs of media are significantly different* – what do you find?