HUDM 5026 - Introduction to Data Analysis and Graphics in R
HW 07 – Tidy Data

Instructions.

- Use R Markdown to create an html document with the homework tasks.
- You are encouraged to discuss problems with classmates, but all work you submit must be your own.
- As always, any plots should have appropriate axis and overall labels.

Data for today's HW come from Maxwell and Delaney's book on experimental design. In particular, we will use data from chapter 11, exercise 19. The data are modeled after an experiment called "Pattern vision in newborn infants" published in *Science* in 1963. From Maxwell & Delaney, p. 571:

> Until the 1960s it was believed that infants had little or no pattern vision during the early weeks or even months of their lives. Fourteen infants under 48 hrs old were exposed to a series of targets, presented in a random sequence to each infant. Three of the targets contained black-and-white patterns: a schematic face, concentric circles, and a section of newspaper. The fourth target was an unpatterned white circle. A blue background was provided in all cases to contrast with the target. The dependent measure is the length of gaze (in sec) of an infant at a particular target.

To get the data in R, install and load package **AMCP**, which is associated with Maxwell and Delaney's book. To access particular data files from the package, use the `data()` function. To get the data set from chapter 11, exercise 19, run `data("C11E19")`. You should then see the data set in your environment.

1. Rename the data set `dat1`.
2. Each row of `dat1` represents data from a different participant. Add a participant ID number column to the data set and call it `ID`.
3. Create `dat1_long` by pivoting `dat1` from wide to long. When you pivot, name the new column that will hold the categorical values (Face, Circle, Newspaper, and White) `target` and name the new column that will hold the numeric outcome data `gaze_time`.
4. Use `apply()` with `dat1` to find the sample for each child for each target (Face, Circle, Newspaper, and White).
5. Use `summarize()` and `group_by()` with `dat1_long` to find the sample mean for each child for each target (Face, Circle, Newspaper, and White).
6. Use `ggplot()` with `dat_long` to create a conditional boxplot that has separate boxplots of gaze time for each target (Face, Circle, Newspaper, and White).