Homework 1

The homework is in this format to facilitate you showing your code and output. Render the document when you are finished and submit it on Canvas. If you have trouble rendering to Word, try format: html instead. If you are still having trouble, you can always copy and paste your code and output (or screenshots of the output) into a Word document. What is important is that we can see both your code and its output.

1. Start a new R session. Before you can read in the data, load the tidyverse set of packages. Include your code to do so. Then read in the data.
2. Use mutate() to create a new variable that is income per child in the respondent’s household growing up, income\_per\_child. The number of siblings is in the nsib variable, and the respondent also counts as a child in the household. (Note: this and any of the other questions can take multiple steps if you need them – you don’t have to do it all in one line of code. You don’t have to make a different dataset every time, but can keep storing everything as nlsy <- mutate(nlsy, ...) for simplicity.)
3. What is the mean income per child in the household? What is the *median* income per child in the household?
4. Create a factor variable (sibling) that has the value “sister”, “brother”, or “only child” depending on a person’s sex and number of siblings. E.g. if the individual has siblings and is male, they are a “brother” to their siblings.
5. How many sisters, brothers, and only children are in the dataset?
6. Redefine the sibling variable so that the categories are ordered by size, with the smallest as the reference category.
7. The variable age\_bir refers to the age of a respondent when they had their first child. Use case\_when() to create the following age categories: “teenager” (13-19), “young adult” (20-29), “adult” (30+). Check to sure it is a correctly labeled factor variable in chronological order (show the code you use to do so).