Homework 1

Please copy and paste your code and answers into this document and submit on Canvas.

1. Using the NLSY data, make a scatter plot of the relationship between hours of sleep on weekends and weekdays. Color it according to region (where 1 = northeast, 2 = north central, 3 = south, and 4 = west).
2. Replace geom\_point() with geom\_jitter(). What does this do? Why might this be a good choice for this graph? Play with the width = and height = options. This site may help: https://ggplot2.tidyverse.org/reference/geom\_jitter.html
3. Use the shape = argument to map the sex variable to different shapes. Change the shapes to squares and diamonds. (Hint: how did we manually change colors to certain values? This might help: https://ggplot2.tidyverse.org/articles/ggplot2-specs.html)
4. When we’re comparing distributions with very different numbers of observations, instead of scaling the y-axis like we did with the facet\_grid() function, we might want to make density histograms. Use google to figure out how to make a density histogram of income. Facet it by region.
5. Make each of the regions in your histogram from part 1 a different color. (Hint: compare what col = and fill = do to histograms).
6. Instead of a log-transformed x-axis, make a square-root transformed x-axis.
7. Doing part 3 squishes the labels on the x-axis. Using the breaks = argument that all the scale\_x\_() functions have, make labels at 1000, 10000, 25000, and 50000.