

HUDM 5026 - Introduction to Data Analysis and Graphics in R

HW 03 – Data Frames and Bivariate Plots

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.
1. There is a data set in R called `mtcars`.
 - (a) Is `mtcars` a data frame or a matrix? If it is not a data frame, make it one using `data.frame()`.
 - (b) Access the help file on `mtcars` and summarize the brief description of the data set.
 - (c) Of the variables in `mtcars` all are of class numeric. That said, are there any that it would make sense to call categorical? Which ones and why?
 2. Create a scatterplot of the relationship between horsepower (horizontal axis) and miles per gallon (vertical axis). Add a linear regression line to the plot and describe the relationship in words in context. Do this in both base R and **ggplot2**.
 3. Make a conditional violin plot with superimposed boxplots in **ggplot2** of miles per gallon conditional on transition type (automatic or not). Interpret the plot in context.
 4. Use `factor()` to add factor versions of the variables `gear` and `cyl`. Create dodged and stacked bar charts in **ggplot2**. For your stacked bar charts, look up how to normalize the bars so that they equal 100% and do that. Interpret the results in context.
 5. Use `mtcars` data to create a scatterplot that pulls in multivariate information like we did with the state abbreviations and regions in class. Surprise me.