1. Using the mtcars dataset, run a two-way ANOVA to test for effects of auto/manual transmission (the ‘am’ column), cylinder (the ‘cyl’ column), and the interaction between the two on mpg.
   1. Is mpg normally distributed?
   2. Run the two-way ANOVA, interpret the results (make sure to include interpretation for the interaction)
   3. Make a boxplot to visualize your results
   4. For this question, because we are doing an ANOVA the ‘am’ variable and ‘cyl’ variable need to be factors. Make sure to run the following:
      1. mtcars<-as.data.frame(mtcars)
      2. mtcars$am<-as.factor(mtcars$am)
      3. mtcars$cyl<-as.factor(mtcars$cyl)
2. Using the mtcars dataset, run an ANOVA to test for the effect of number of cylinders (the ‘cyl’ column) on mpg.
   1. Is mpg normally distributed?
   2. Run the ANOVA, interpret the results
   3. Run the appropriate post-hoc test to look for specific differences between all pairs of cylinders, interpret the results
   4. Make a boxplot to visualize your results
   5. For this question, because we are doing an ANOVA the ‘cyl’ variable needs to be a factor. Make sure to run the following:
      1. mtcars<-as.data.frame(mtcars)
      2. mtcars$cyl<-as.factor(mtcars$cyl)
3. Using the data about admission to graduate school, look for differences in means of GPAs of students at schools of different ranks.
   1. You need the following code to access this dataset:
      1. install.packages("aod")
      2. library(aod)
      3. admit.data <- read.csv("https://stats.idre.ucla.edu/stat/data/binary.csv")
   2. Transform the GPA data to look as normal as possible, run an ANOVA
   3. Run the non-parametric ANOVA on non-transformed data, are the results different from the ANOVA?
   4. Make a boxplot to visualize the results (use the non-transformed data for this)
   5. For this question, because we are doing an ANOVA, rank must be a factor. Make sure to run the following:
      1. admit.data$rank<-as.factor(admit.data$rank)
4. Using the data about admission to graduate school, look for a correlation between GPA and GRE
   1. You need the following code to access this dataset:
      1. install.packages("aod")
      2. library(aod)
      3. admit.data <- read.csv("https://stats.idre.ucla.edu/stat/data/binary.csv")
   2. Make a scatter plot with GPA on the x-axis and GRE on the y-axis
   3. Are GPA and GRE scores normally distributed?
   4. Run a correlation test (make sure it’s the correct one for the data)
   5. Interpret p and r values
5. Using the data about admission to graduate school, make a linear model that predicts values of GRE scores based on GPA values. You may break some assumptions of linear models here if the data require you to do so.
   1. You need the following code to access this dataset:
      1. install.packages("aod")
      2. library(aod)
      3. admit.data <- read.csv("https://stats.idre.ucla.edu/stat/data/binary.csv")
   2. Run a linear model, interpret the summary table
   3. Make a scatter plot with GPA on the x-axis and GRE on the y-axis, add the best fit line
   4. Look at the diagnostic plots, tell me what you conclude from them
6. Using the mtcars dataset, make a linear model that predicts the mpg values of cars based on their weight (wt column)
   1. Make a scatter plot with weight on the x-axis and mpg on the y-axis
   2. Run a linear model, interpret the summary table
   3. Look at the diagnostic plots, tell me what you conclude from them
   4. Predict 10 values of your choice that aren’t included in the data
7. Design an experiment that requires one of the following statistical analyses: two-way ANOVA, correlation, linear regression, or logistic regression. What is your biological question? What is your hypothesis? What are you independent and dependent variables? Are they numeric, categorical, or ranked? Which test will you be using? What is the null hypothesis of the test? How will you analyze non-normal data if your results happen to be non-normal?