Section Assignment 9

Please Read!

Does learning a second language change brain structure? Machelli et al.(2004)* tested 22 native Italian speakers who had learned English as a second language. Proficiencies in reading, writing, and speech were assessed by a proficiency score. Gray matter density was measured in the left inferior parietal region of the brain using a neuroimaging technique, as mm^3 of gray matter per voxel.

- Machelli et al.(2004) Structural plasticity in the bilingual brain. Nature 431:757 https://doi.org/10.1038/431757a

You will find the data from their study in the "Week9Brains.csv" file on Canvas.

Since we have already downloaded and installed the tidyverse, this week, simply load the tidyverse functions into your working library:

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.6     v dplyr   1.0.7
## v tidyr   1.1.4     v stringr 1.4.0
## v readr   2.1.1     v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

This line of code above will need to be run each time you start a new R session.

One of the traits of a great programmer is their ability to solve a problem they haven't seen before. One of the best ways to solve a problem you have not seen before is to see if anyone else has. Great programmers are experts of web searching. Do not be afraid to find code on an online help page and run it!

Copying from the internet is one of the foundations of learning how to program, but it only works as learning if you reflect on why the code you used works. For that reason, this week and in future weeks, you will need to annotate your code or answer a specific question about the process. Use the lines that look like this at the end of each question to input your explanation of the code:

```
note<-'your note here'
print(note)
```

You will be graded on not just your code, but your explanation. Remember, in this class, it is okay to copy code, but you still need to demonstrate independent thought.

Just like last week, this week you will practice answering a research question step by step. Think of it as practice for your term project!

Step 1. Import your data, and check the assumption of normality.

Question 1. (0.5 point): Import the data you downloaded from canvas this week.

```
BrainData<-read.csv("Week9Brains.csv")
head(BrainData)
```

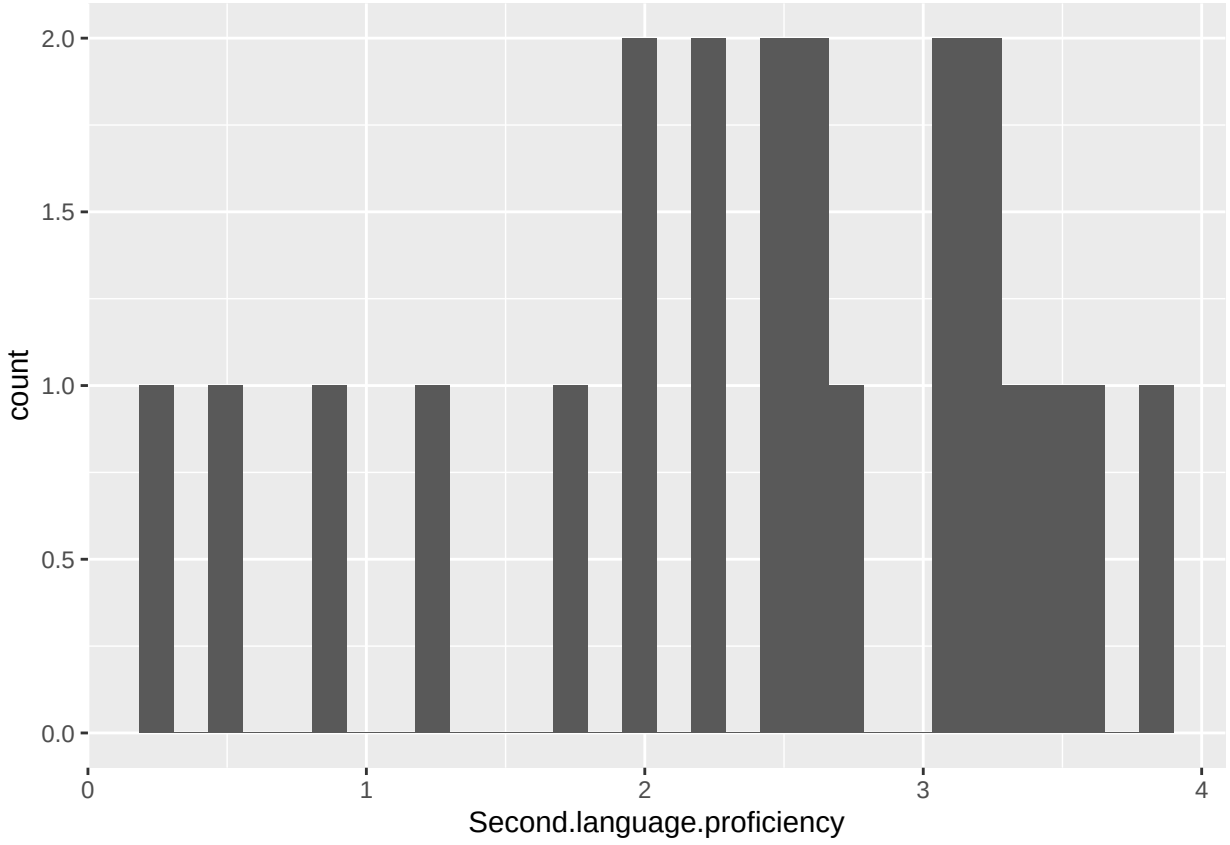
```
##   Second.language.proficiency Gray.matter.density
## 1                        0.26              -0.070
## 2                        0.44              -0.080
## 3                        0.89              -0.008
## 4                        1.26              -0.009
## 5                        1.69              -0.023
## 6                        1.97              -0.009
```

Question 2. (2 points): Use visualizations and tests to determine that the data are normally distributed. Yes, you will have to do this twice.

```
#visualize the second language scores data here.
Histogram1<-ggplot(data = BrainData,mapping = aes(x = Second.language.proficiency))+
  geom_histogram()

Histogram1
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
#Test for normality of the second language scores in this code box.
ks.test(x = BrainData$Second.language.proficiency,pnorm,mean(BrainData$Second.language.proficiency),sd(
```

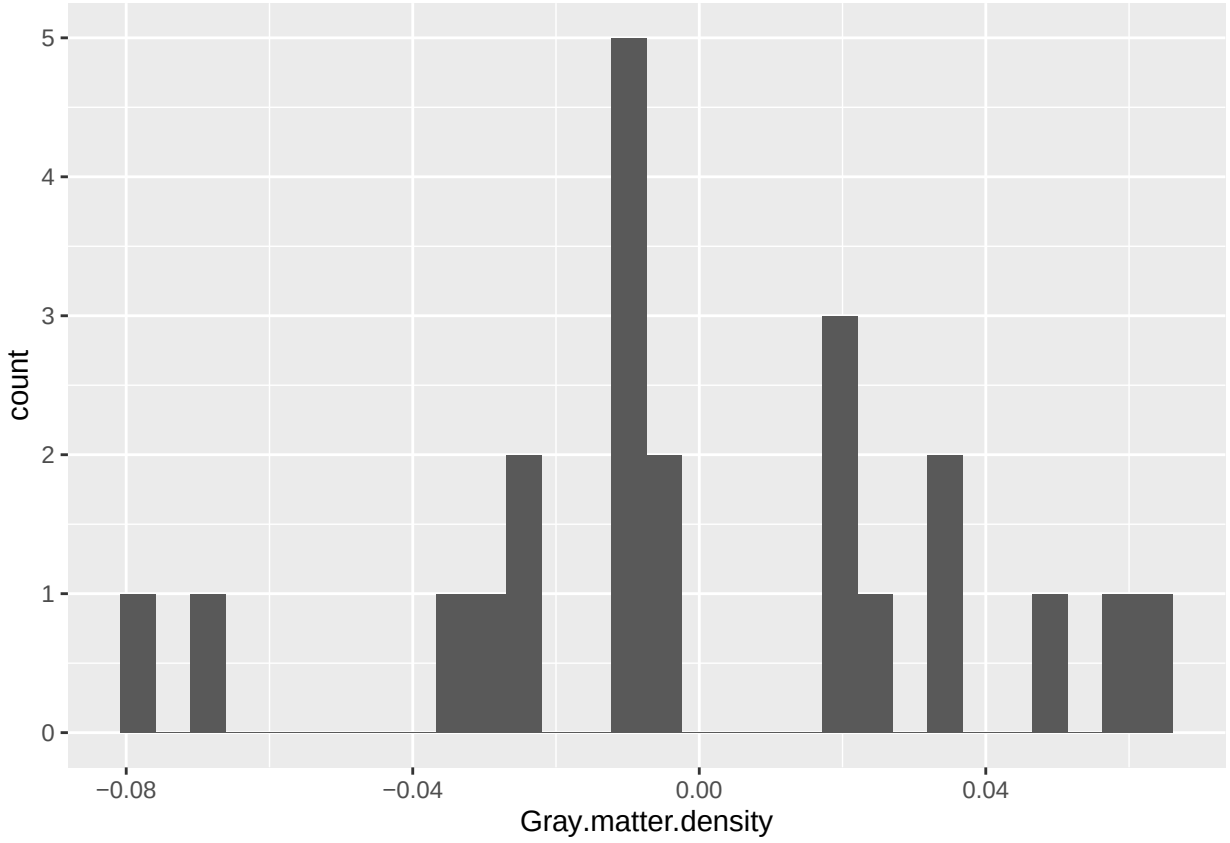
```
## Warning in ks.test(x = BrainData$Second.language.proficiency, pnorm,
## mean(BrainData$Second.language.proficiency), : ties should not be present for
## the Kolmogorov-Smirnov test
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  BrainData$Second.language.proficiency
## D = 0.13079, p-value = 0.846
## alternative hypothesis: two-sided
```

Are the second language scores data normally distributed? Is it okay that the histogram looks weird? Why/why not? [1] ""

```r
#visualize the gray matter density data here.
Histogram2<-ggplot(data = BrainData,mapping = aes(x = Gray.matter.density))+
  geom_histogram()
Histogram2
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```r
#Test for normality of the gray matter density data in this code box.
ks.test(x = BrainData$Gray.matter.density,pnorm,mean(BrainData$Gray.matter.density),sd(BrainData$Gray.ma
```

```
## Warning in ks.test(x = BrainData$Gray.matter.density, pnorm,
## mean(BrainData$Gray.matter.density), : ties should not be present for the
## Kolmogorov-Smirnov test
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  BrainData$Gray.matter.density
## D = 0.15955, p-value = 0.6299
## alternative hypothesis: two-sided
```

Are the gray matter density data normally distributed? Is it okay that the histogram looks weird? Why/why not? [1] ""

Step 2. Identify your null and alternate hypotheses, generate the correct linear model, test the null hypothesis.

Research question: Does learning a second language change brain structure?

Question 3. (0.5 point): What is the null and alternate hypothesis that answer the research question?

[1] "" [1] ""

Question 4. (3 points): Generate the linear model and test your null hypothesis. Consult the PDF help file for code associated with performing statistical hypothesis testing in R.

```
modelWeek9<-lm(BrainData$Gray.matter.density~BrainData$Second.language.proficiency)
summary(modelWeek9)
```

```
##
## Call:
## lm(formula = BrainData$Gray.matter.density ~ BrainData$Second.language.proficiency)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.032123 -0.020623 -0.001865  0.017059  0.037232
##
## Coefficients:
##                                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)                          -0.072142   0.012309  -5.861 9.84e-06 ***
## BrainData$Second.language.proficiency  0.030235   0.004749   6.367 3.26e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02183 on 20 degrees of freedom
## Multiple R-squared:  0.6696, Adjusted R-squared:  0.6531
## F-statistic: 40.54 on 1 and 20 DF,  p-value: 3.264e-06
```

Interpret the results of the test you performed to answer the research question in the note below. Does learning a second language change brain structure? Hint:Your answer should include information you gained from the p-value, the slope, and the R-squared.

[1] ""

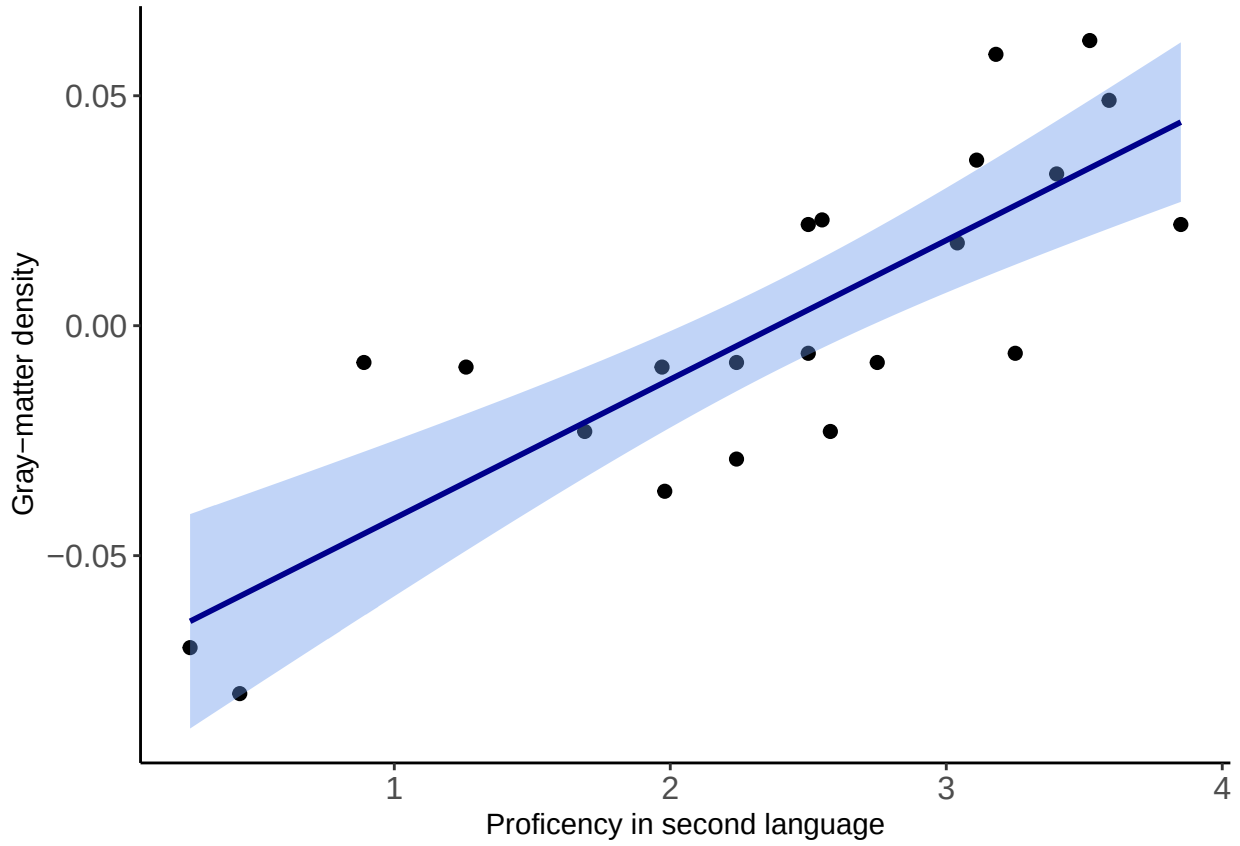
Step 3. Graphical representation of the model.

Question 5. (2 points): Generate a figure representing the data, your model and the error around your model. Use my example as a template. Drawing a line of best fit can be done a bunch of ways in ggplot, but I recommend looking into how to use the `geom_smooth()` function as it can also create the shaded error on the graph, which is required to get full credit. Don't worry about matching the colors to the example, in fact I encourage you to look up all the different color options in R and find a color scheme you like! Some hints to help you: 1) The size of the points is larger than the default. 2) The axis text for both x and y is increased. 3) The x and y axis need to be relabeled.

```
plot1<-ggplot(data = BrainData,mapping = aes(x=Second.language.proficiency,y = Gray.matter.density))+
  geom_point(size=2)+
  geom_smooth(method = "lm",color="darkblue",fill="cornflowerblue")+
  theme_classic()+
  theme(axis.text.x = element_text(size=12))+
  theme(axis.text.y = element_text(size=12))+
  xlab(label = "Proficency in second language")+
  ylab(label = "Gray-matter density")

plot1
```

```
## `geom_smooth()` using formula 'y ~ x'
```



This graph is definitely competing with last week's for most complex one we have made yet! Explain how you got the model line and error on the graph. [1] ""

Part 2: A question on correlation: In the scenario above, the scientists were confident that they had an independent and dependent variable. What if we did not share that opinion? Reanalyze the brain data using a correlative model instead.

Step 1: We will assume the data are normally distributed, you already tested this in Q2.

Step 2. Identify your null and alternate hypotheses, generate the correct model, test the null hypothesis.

Research question: Is there a relationship between learning a second language and brain structure?

Question 6. (0.5 point): What is the null and alternate hypothesis that answer the research question?

[1] "" [1] ""

Question 7. (1.5 points): Test the null hypothesis. Consult the PDF help file for code associated with performing statistical hypothesis testing in R.

```
cor.test(BrainData$Second.language.proficiency,BrainData$Gray.matter.density)
```

```
##
##  Pearson's product-moment correlation
##
## data:  BrainData$Second.language.proficiency and BrainData$Gray.matter.density
## t = 6.3671, df = 20, p-value = 3.264e-06
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.6056638 0.9218696
## sample estimates:
##       cor
## 0.8183134
```

Interpret the results of the test you performed to answer the research question in the note below. Is there a relationship between learning a second language and brain structure? Hint: Your answer should include information you gained from the p-value, the correlation, and the confidence interval.

[1] ""

Once you are done, click the “Knit” button above(It looks like a blue ball of yarn). Save the pdf file with your name and the week number in the file name:

(for example: “Liam_Mueller_Section_9.pdf”).

Then upload this pdf file to canvas/gradescope under the Week 9 Section Assignment before the deadline.