prelim

Hewlett Madrid

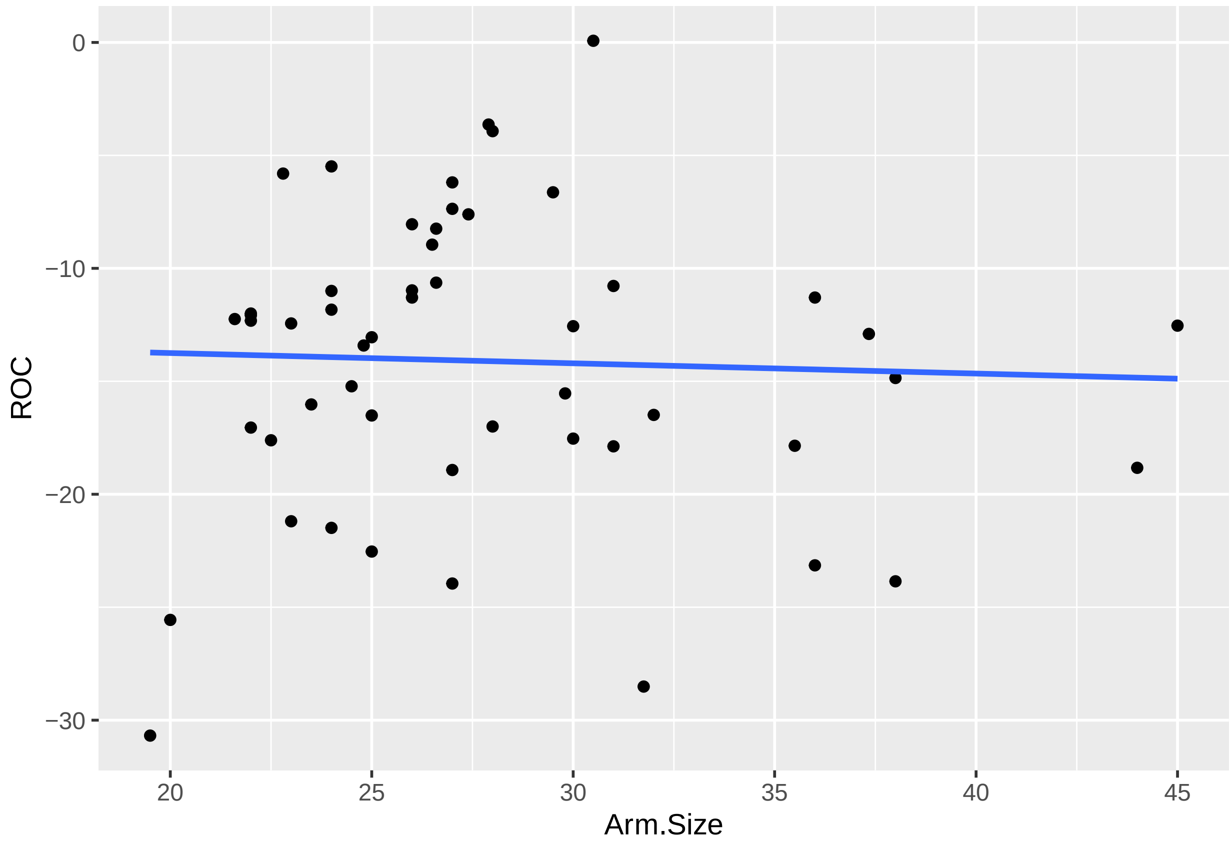
2023-04-20

```
library(tidyverse)
library(ggpubr)
library(broom)
library(car)
library(rstatix)
```

```
roc <- read.csv("C:/Users/Xyrine/Documents/School Stuff/BS BIO 4th Year/1st Semester/BIO 118/Module 1/Mo

colnames(roc) <- c("Arm.Size", "ROC")

ggplot(roc, aes(Arm.Size, ROC)) + geom_point() +
  geom_smooth(method = "lm", se = FALSE)
```



```r
roc.lm <- lm(ROC ~ Arm.Size, roc)

summary(roc.lm)
```

```
##
## Call:
## lm(formula = ROC ~ Arm.Size, data = roc)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -16.956  -3.458   1.485   3.428  14.299
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -12.84094    4.55287  -2.820  0.00686 **
## Arm.Size     -0.04541    0.16037  -0.283  0.77824
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.549 on 50 degrees of freedom
## Multiple R-squared:  0.001601,   Adjusted R-squared:  -0.01837
## F-statistic: 0.08017 on 1 and 50 DF,  p-value: 0.7782
```

```r
cor.test(roc$Arm.Size, roc$ROC)
```

```
##
##  Pearson's product-moment correlation
##
## data:  roc$Arm.Size and roc$ROC
## t = -0.28314, df = 50, p-value = 0.7782
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.3095303  0.2354616
## sample estimates:
##         cor
## -0.04000959
```

```r
#  imports pooled data into R
msc <- read.csv("C:/Users/Xyrine/Documents/School Stuff/BS BIO 4th Year/1st Semester/BIO 118/Module 1/Mo


#  isolates columns for arm size and lengths, and velocity
msc.cont <- msc %>%
  select(Upper.Arm.Size, Upper.Arm.Length, Forearm.Length,
         X0.25, X0.5, X1, X1.5, X2) %>%
  drop_na() %>%
  gather(Load, Velocity, X0.25, X0.5, X1, X1.5, X2)
```

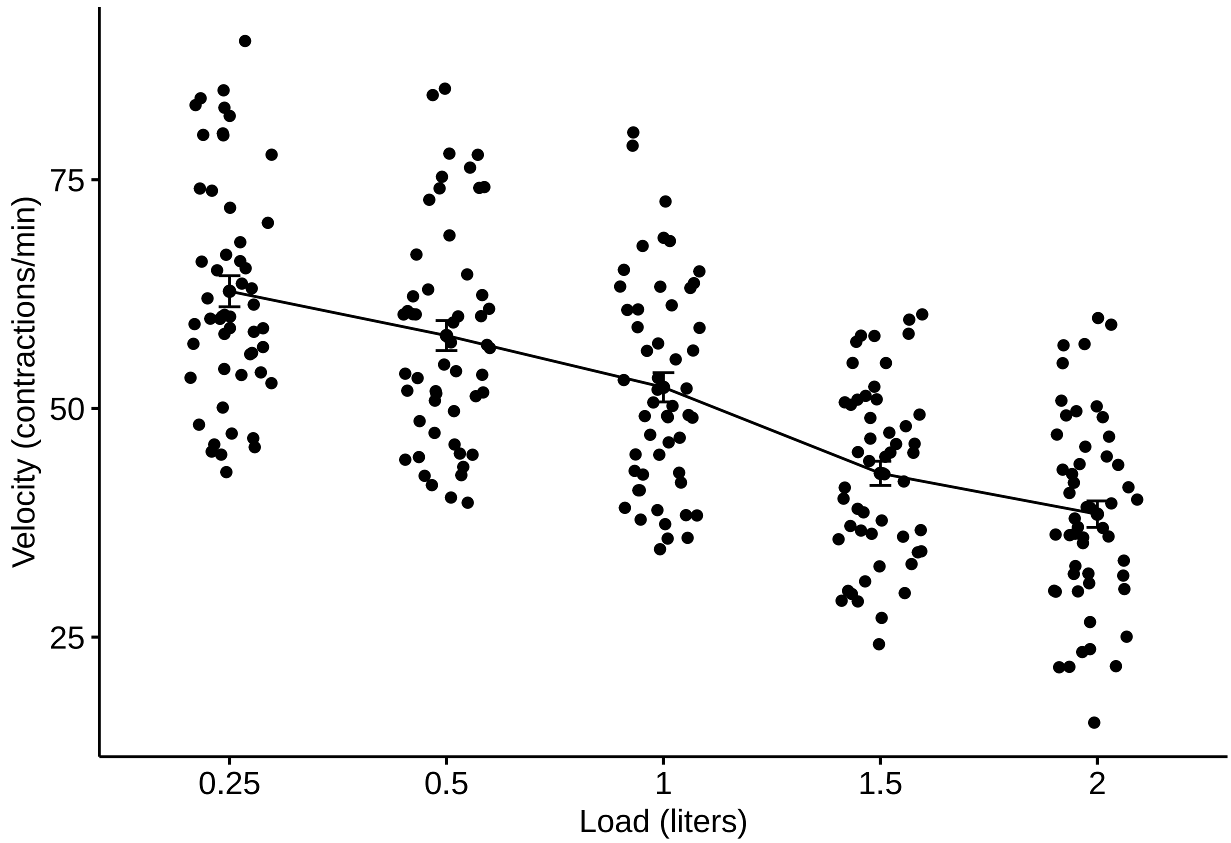
1. Load and Velocity

1.a. Data summary and visualization

```
#  prints summary of data
msc.cont %>%
  group_by(Load) %>%
  summarise(
    Mean.Velocity = mean(Velocity),
    SD.Velocity = sd(Velocity)
  )
```

```
## # A tibble: 5 x 3
##   Load  Mean.Velocity SD.Velocity
##   <chr>         <dbl>       <dbl>
## 1 X0.25          62.8       12.2
## 2 X0.5           58.0       11.8
## 3 X1             52.3       11.5
## 4 X1.5           42.9        9.53
## 5 X2             38.4       10.4
```

```
#  visualizes data
ggline(msc.cont, x = "Load", y = "Velocity",
       add = c("mean_se", "jitter"),
       order = c("X0.25", "X0.5", "X1", "X1.5", "X2"),
       ylab = "Velocity (contractions/min)", xlab = "Load (liters)") +
   scale_x_discrete(labels=c("0.25","0.5", "1", "1.5", "2"))
```



1.b. ANOVA

```r
#  Assumption: homogeneity of variances
leveneTest(Velocity ~ Load, data = msc.cont)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##        Df F value Pr(>F)
## group   4  0.6653 0.6166
##       255
```

```r
## Accept null hypothesis (p > 0.05)
## Not enough evidence suggests that differences in variances between groups
### are statistically significant

## Assume homoscedasticity


# Assumption: normality
shapiro.test(msc.cont$Velocity[msc.cont$Load == "X0.25"])  #  (p < 0.05)
```

```
##
##  Shapiro-Wilk normality test
##
```

```
## data:  msc.cont$Velocity[msc.cont$Load == "X0.25"]
## W = 0.95087, p-value = 0.03171
```

```r
shapiro.test(msc.cont$Velocity[msc.cont$Load == "X0.5"])   #  (p < 0.05)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  msc.cont$Velocity[msc.cont$Load == "X0.5"]
## W = 0.95079, p-value = 0.03147
```

```r
shapiro.test(msc.cont$Velocity[msc.cont$Load == "X1"])     #  (p > 0.05)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  msc.cont$Velocity[msc.cont$Load == "X1"]
## W = 0.95969, p-value = 0.07586
```

```r
shapiro.test(msc.cont$Velocity[msc.cont$Load == "X1.5"])   #  (p > 0.05)
```

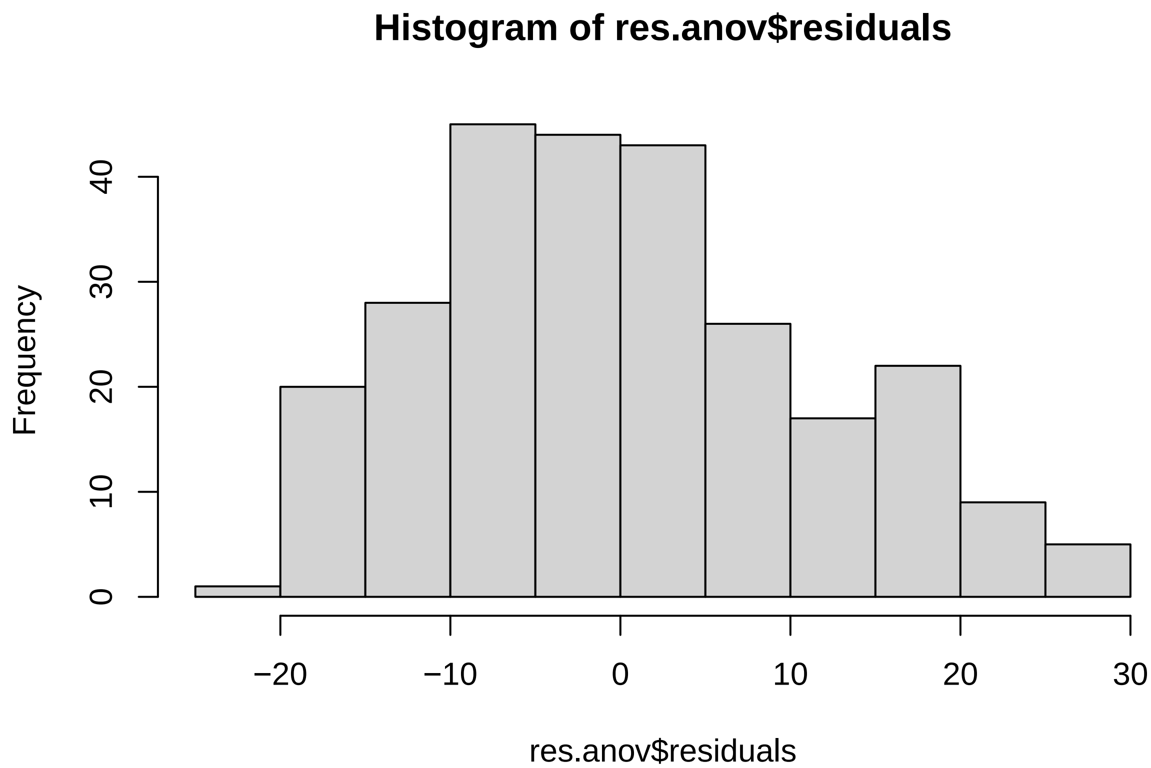
```
##
##  Shapiro-Wilk normality test
##
## data:  msc.cont$Velocity[msc.cont$Load == "X1.5"]
## W = 0.97211, p-value = 0.2591
```

```r
shapiro.test(msc.cont$Velocity[msc.cont$Load == "X2"])     #  (p > 0.05)
```

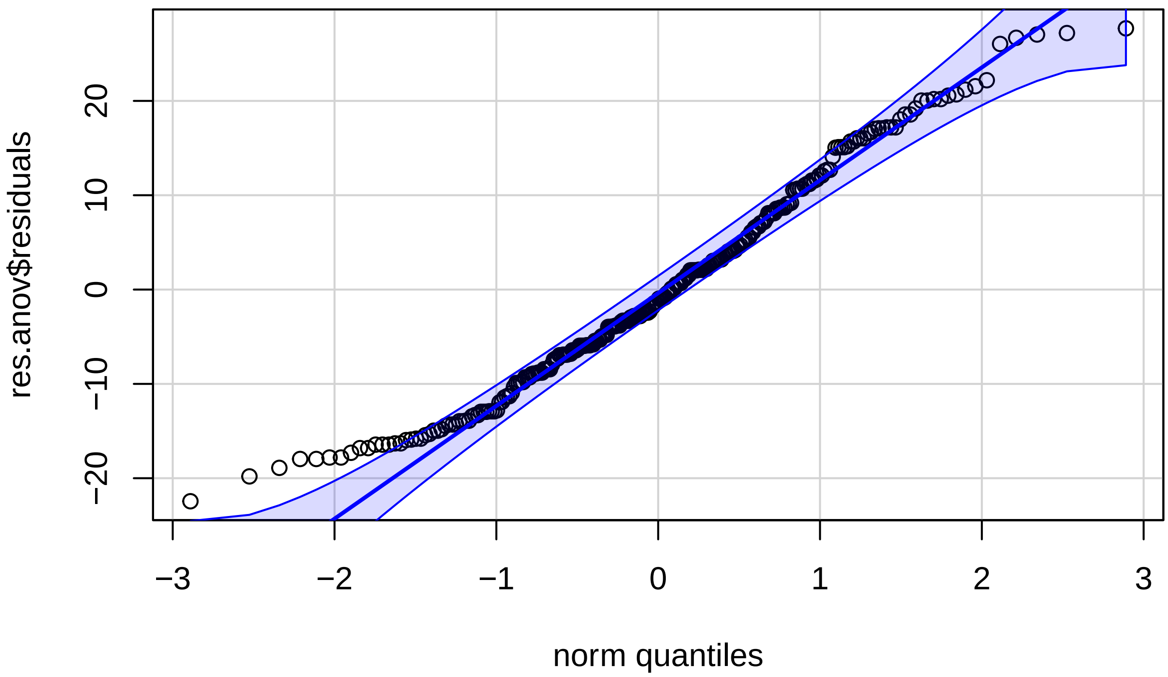
```
##
##  Shapiro-Wilk normality test
##
## data:  msc.cont$Velocity[msc.cont$Load == "X2"]
## W = 0.98488, p-value = 0.7461
```

```r
res.anov <- aov(Velocity ~ Load, msc.cont)

## histogram
hist(res.anov$residuals)
```




```r
## qqplot
qqPlot(res.anov$residuals,
  id = FALSE # id = FALSE to remove point identification
)
```




```
## histogram shows roughly a bell curve, and points in the QQ plot nearly adhere to
### a streight line and are well within the confidence bands

## We can assume normality


#  ANOVA
anov <- aov(Velocity ~ Load, msc.cont)
summary(anov)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## Load          4  21464    5366   43.14 <2e-16 ***
## Residuals   255  31718     124
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```