

# LAB4

2025-04-26

## R Markdown

### 1a

Do this using the aov function (i.e. you do not need to perform calculations from scratch here, nor do you have to do a permutation test). Be sure to state your null and alternative hypotheses, and report both your statistical conclusion (reject / fail to reject  $H_0$ ) and a statement in the context of the problem.

```
# Enter the data for each technique
mt1 <- c(3129, 3000, 2865, 2890)
mt2 <- c(3200, 3300, 2975, 3150)
mt3 <- c(2800, 2900, 2985, 3050)
mt4 <- c(2600, 2700, 2600, 2765)

# Create data frame
df_w <- data.frame(strength = c(mt1, mt2, mt3, mt4),
                    mt = c(
                      rep("1", length(mt1)),
                      rep("2", length(mt2)),
                      rep("3", length(mt3)),
                      rep("4", length(mt4))
                    ))
sample(df_w) #Print out the sample of the data.frame
```

```
##      strength mt
## 1      3129  1
## 2      3000  1
## 3      2865  1
## 4      2890  1
## 5      3200  2
## 6      3300  2
## 7      2975  2
## 8      3150  2
## 9      2800  3
## 10     2900  3
## 11     2985  3
## 12     3050  3
## 13     2600  4
## 14     2700  4
## 15     2600  4
## 16     2765  4
```

```
#Fit a one-way ANOVA model to test if mean tensile strength differs across mixing techniques
#'Fvalue' stores the ANOVA summary table, which includes the F statistic and p-value for the test.
Fvalue_1 <- summary(aov(strength ~ mt, data = df_w))
Fvalue_1
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## mt           3 489740   163247    12.73 0.000489 ***
## Residuals    12 153908    12826
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Report the null

\$ H\_0\$: \$ All group means are equal

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

## Report the alternative hypotheses

\$ H\_a\$: \$ At least one of these groups has different mean.

## Report both your statistical conclusion (reject / fail to reject H0)

F value: 12.73 p-value: 0.000489 The p-value is less than  $\alpha = 0.05$ , we reject the null hypothesis \$ (H\_0)\$ \$

The method of mixing Portland cement significantly affects its tensile strength. In other words, at least one of the four mixing techniques produces a mean tensile strength that is different from the others.

## 1b

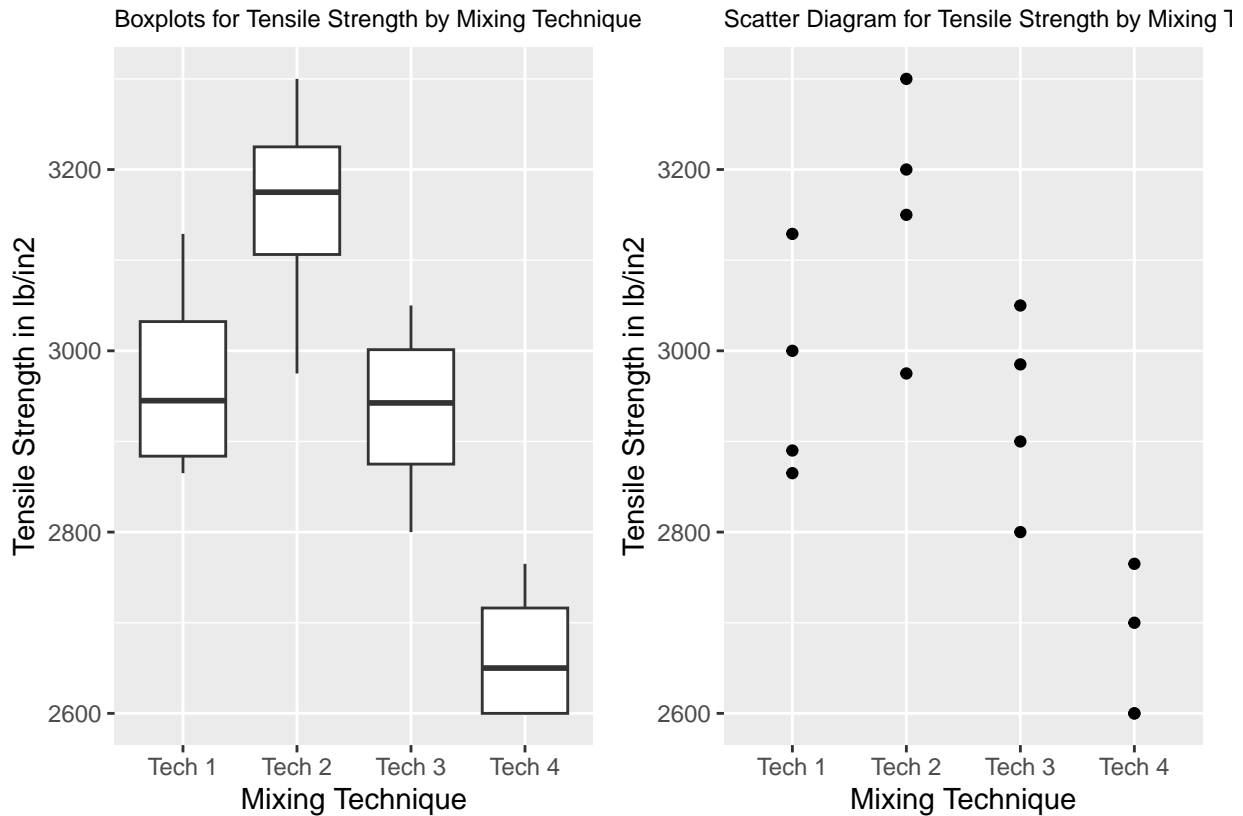
Instead, make the two types of graphs we discussed in Lecture #5 for these type of data (side-by-side boxplots and a scatter diagram). Briefly describe what you observe.

```
library(ggplot2)
library(patchwork)
#Boxplot
p1 <- ggplot() +
  geom_boxplot(aes(x = "Tech 1", y = mt1)) +
  geom_boxplot(aes(x = "Tech 2", y = mt2)) +
  geom_boxplot(aes(x = "Tech 3", y = mt3)) +
  geom_boxplot(aes(x = "Tech 4", y = mt4)) +
  scale_x_discrete(limits=c("Tech 1", "Tech 2", "Tech 3", "Tech 4")) +
  ggtitle("Boxplots for Tensile Strength by Mixing Technique") +
  ylab("Tensile Strength in lb/in2") +
  xlab("Mixing Technique") +
  theme(plot.title = element_text(size = 8.5))

#Scatter plot
p2 <- ggplot() +
  geom_point(aes(x = "Tech 1", y = mt1)) +
  geom_point(aes(x = "Tech 2", y = mt2)) +
```

```
geom_point(aes(x = "Tech 3", y = mt3)) +
geom_point(aes(x = "Tech 4", y = mt4)) +
scale_x_discrete(limits=c("Tech 1", "Tech 2", "Tech 3", "Tech 4")) +
ggtitle("Scatter Diagram for Tensile Strength by Mixing Technique") +
ylab("Tensile Strength in lb/in2") +
xlab("Mixing Technique") +
theme(plot.title = element_text(size = 8.5))
```

p1 + p2



Description: The boxplot and scatter plot both show that Mixing Tech 2 results in the highest tensile strengths. And then followed by Tech 1 and Tech 3, with Tech 4 has the lowest values. Tech 2 shows more spread between the lowest point and the highest point.

## 1c

Instead of the Fisher LSD method, run the TukeyHSD test to make comparisons between all pairs of means. Report your conclusions at  $\alpha = 0.05$ , including statements in the context of the problem.

```
# Perform Tukey's Honest Significant Difference (HSD) post-hoc test to determine
# Which pairs the mixing techniques that have significantly different mean tensile strengths
modell1 <- aov(strength ~ mt, data = df_w)
tukey_result_1 <- TukeyHSD(modell1)
tukey_result_1
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = strength ~ mt, data = df_w)
##
## $mt
##      diff      lwr      upr    p adj
## 2-1  185.25 -52.50029 423.00029 0.1493561
## 3-1  -37.25 -275.00029 200.50029 0.9652776
## 4-1 -304.75 -542.50029 -66.99971 0.0115923
## 3-2 -222.50 -460.25029  15.25029 0.0693027
## 4-2 -490.00 -727.75029 -252.24971 0.0002622
## 4-3 -267.50 -505.25029 -29.74971 0.0261838
```

Significant differences ( $p < 0.05$ ) 4-1:  $p = 0.0116$  4-2:  $p = 0.0003$  4-3:  $p = 0.0262$

Not significant( $p \geq 0.05$ ) 2-1:  $p = 0.149$  3-1:  $p = 0.965$  3-2:  $p = 0.069$

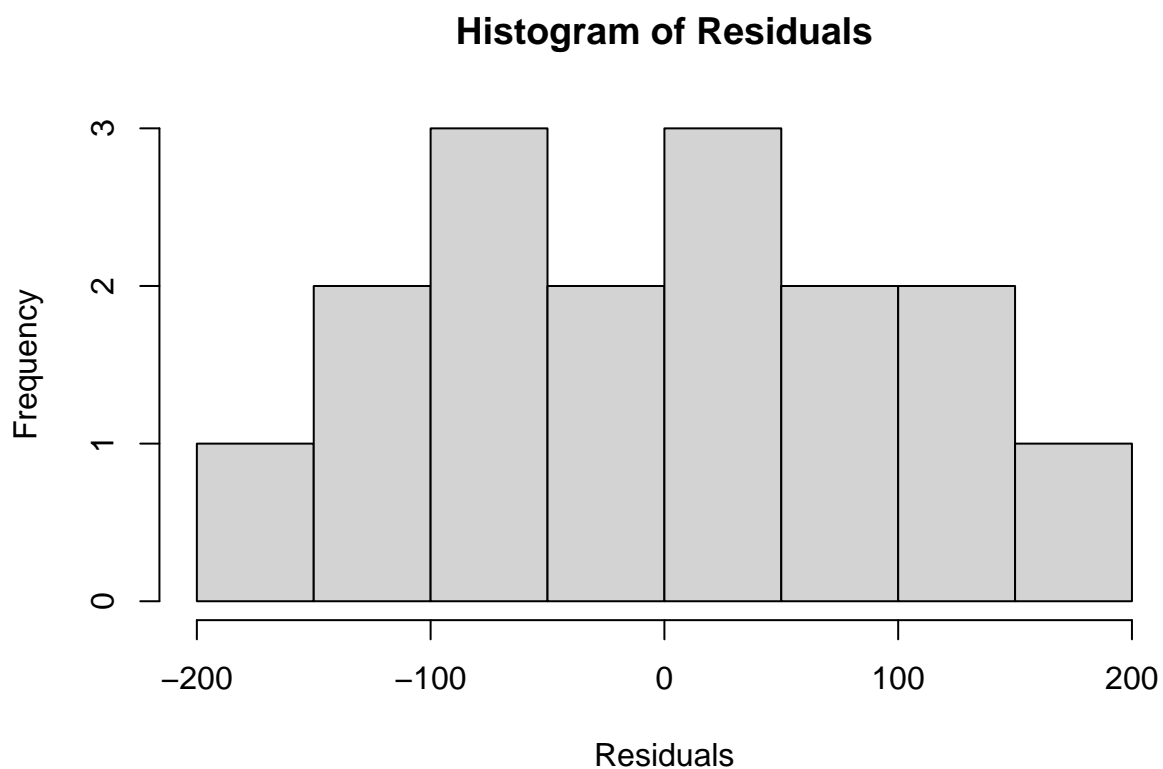
At the  $\alpha = 0.05$  significance level, the TukeyHSD test shows that mixing Technique 4 has a significantly lower mean tensile strength than others. Tech 1,2, and 3 has not much differences between themselves.

## 1d

Construct a normal probability plot as stated, and also a histogram of the residuals and perform the Shapiro-Wilk test. What conclusions do you draw?

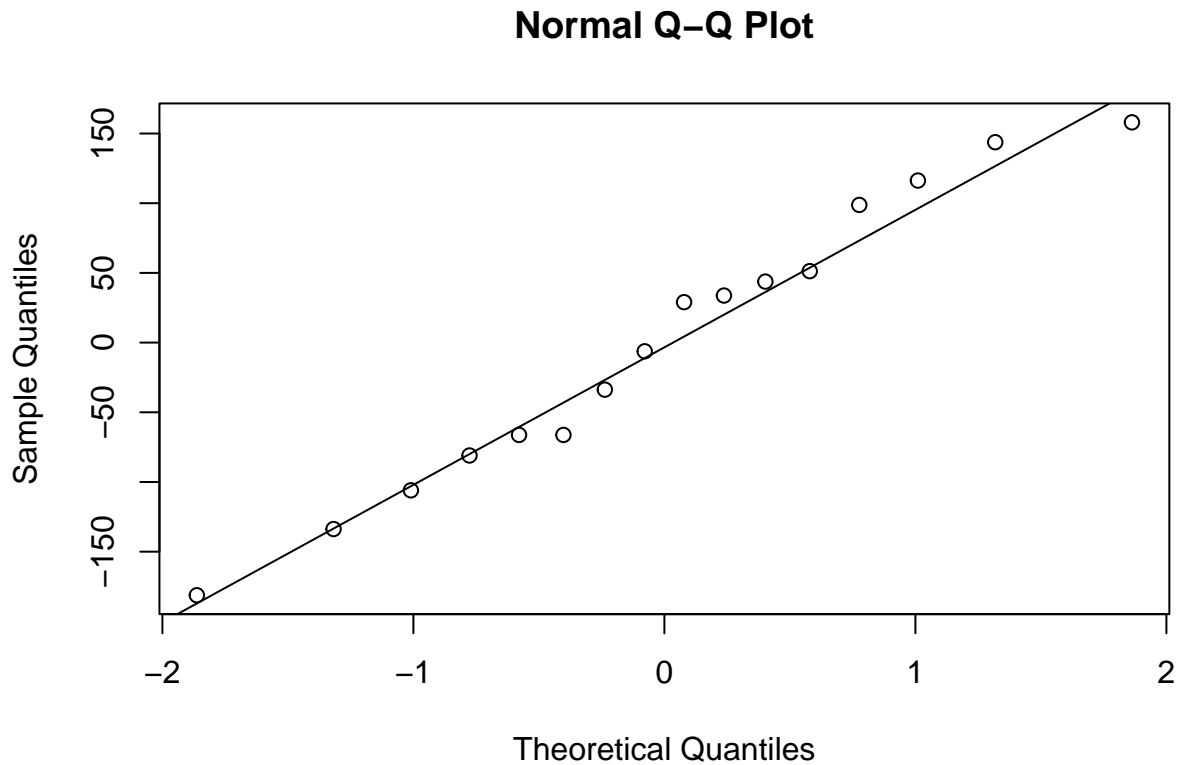
```
# calculate residuals (observation - group mean)
mt1_resid <- mt1 - mean(mt1)
mt2_resid <- mt2 - mean(mt2)
mt3_resid <- mt3 - mean(mt3)
mt4_resid <- mt4 - mean(mt4)

#Combine all residuals
mt_total_resid <- c(mt1_resid, mt2_resid, mt3_resid, mt4_resid)
hist(mt_total_resid, main = "Histogram of Residuals", xlab = "Residuals")
```



Above is the histogram of residual and it appears symmetric which supporting normality.

```
# using the normal probability plot  
qqnorm(mt_total_resid)  
qqline(mt_total_resid)
```



About is the q-q plot and the points fall roughly along the straight line. This means the residuals are approximately normally distributed.

```
#perform the Shapiro Wilk Test
shapiro.test(mt_total_resid)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  mt_total_resid
## W = 0.97046, p-value = 0.846
```

The p-value is 0.846 which is much greater than 0.05. It has fail to reject the  $H_0$ :  $\epsilon$  of normality.

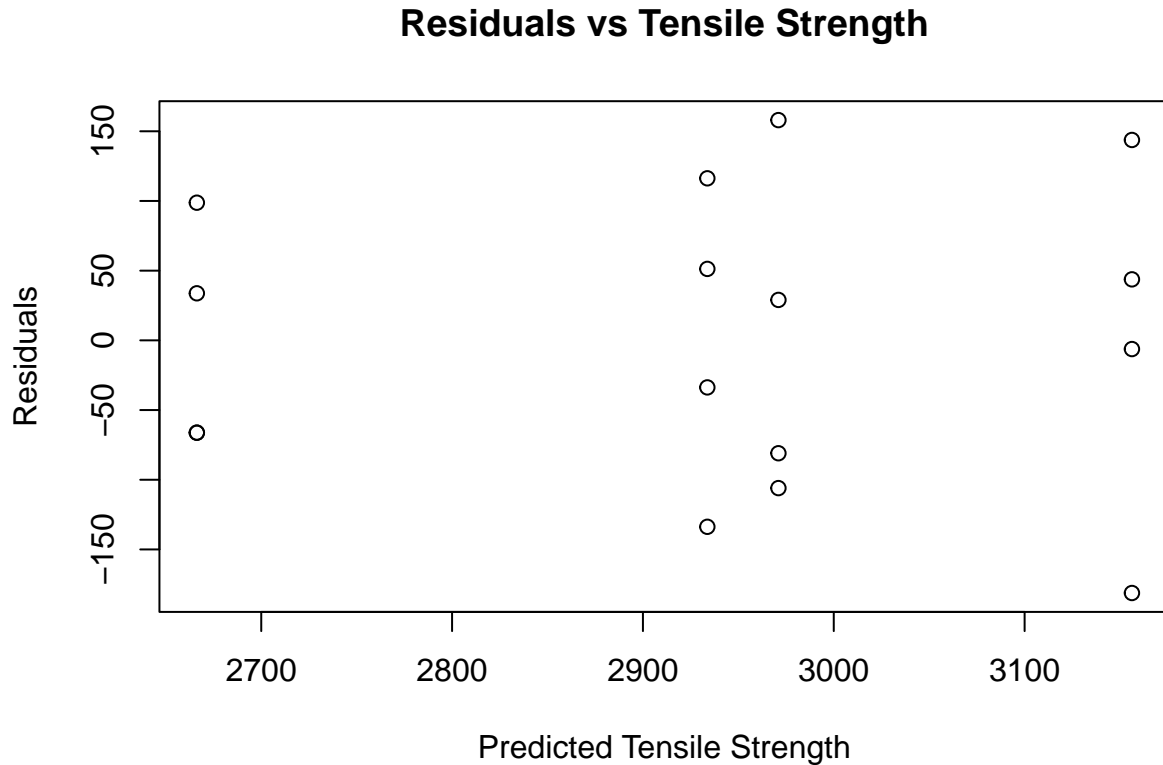
Summary Statement: Based on the Shapiro-Wilk test (p-value = 0.846), the residuals are not normally distributed. Therefore, the normality assumption for ANOVA is satisfied.

**1e**

Plot the residuals versus the predicted tensile strength. Comment on the plot.

```
# fitted (predicted) values and residuals
fitted_tensile_strength1 <- model1$fitted.values
residuals_values1 <- model1$residuals
```

```
# Plot residuals vs the predicted tensile strength
plot(fitted_tensile_strength1, residuals_values1,
     xlab = "Predicted Tensile Strength",
     ylab = "Residuals",
     main = "Residuals vs Tensile Strength")
```



In the residuals vs fitted values plot, the residuals appear randomly scattered around zero with no clear pattern or trend. There are no obvious outliers or systematic structures which supports the validity of the ANOVA model assumptions.

## 2a

Again do this using the aov function (i.e. you do not need to perform calculations from scratch here, nor do you have to do a permutation test, yet). Be sure to state your null and alternative hypotheses, and report both your statistical conclusion (reject / fail to reject  $H_0$ ) and a statement in the context of the problem.

```
# Densities for each temperature group
dens_100 <- c(21.8, 21.9, 21.7, 21.6, 21.7)
dens_125 <- c(21.7, 21.4, 21.5, 21.4)
dens_150 <- c(21.9, 21.8, 21.8, 21.6, 21.5)
dens_175 <- c(21.9, 21.7, 21.8, 21.4)

# Create data frame
df_t <- data.frame(density = c(dens_100, dens_125, dens_150, dens_175),
                  temp = c(rep("100", length(dens_100)),
```

```

rep("125", length(dens_125)),
rep("150", length(dens_150)),
rep("175", length(dens_175)))
sample(df_t)

```

```

##      temp density
## 1    100    21.8
## 2    100    21.9
## 3    100    21.7
## 4    100    21.6
## 5    100    21.7
## 6    125    21.7
## 7    125    21.4
## 8    125    21.5
## 9    125    21.4
## 10   150    21.9
## 11   150    21.8
## 12   150    21.8
## 13   150    21.6
## 14   150    21.5
## 15   175    21.9
## 16   175    21.7
## 17   175    21.8
## 18   175    21.4

```

```

#Fit a one-way ANOVA model to test if mean tensile strength differs across mixing techniques
#'Fvalue' stores the ANOVA summary table, which includes the F statistic and p-value for the test.
Fvalue_2 <- summary(aov(density ~ temp, data = df_t))
Fvalue_2

```

```

##           Df Sum Sq Mean Sq F value Pr(>F)
## temp          3  0.1561  0.05204    2.024  0.157
## Residuals    14  0.3600  0.02571

```

F value = 2.024 p-value = 0.157

Context: The p-value is greater than  $\alpha = 0.05$ , we fail to reject the ( $H_0$ ). The mean densities do not significantly differ across the four temperatures in this experiment.

## Report the null

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

Null hypothesis( $(H_0)$ ): The mean density of bricks is the same at all firing temperatures.

$$H_0 : \mu_{100} = \mu_{125} = \mu_{150} = \mu_{175}$$

## Report the alternative hypotheses Alternative hypothesis ( $H_a$ ): At least one mean density is different.



## 2b

Run the TukeyHSD test to make comparisons between all pairs of means. Report your conclusions at  $\alpha = 0.05$ , including statements in the context of the problem.

```
# Fit the ANOVA model
model2 <- aov(density ~ temp, data = df_t)

# Run the Tukey HSD post-hoc test
tukey_result_2 <- TukeyHSD(model2)
tukey_result_2

##    Tukey multiple comparisons of means
##      95% family-wise confidence level
##
## Fit: aov(formula = density ~ temp, data = df_t)
##
## $temp
##           diff             lwr             upr      p adj
## 125-100 -0.24 -0.55266107  0.07266107  0.1626312
## 150-100 -0.02 -0.31477969  0.27477969  0.9971517
## 175-100 -0.04 -0.35266107  0.27266107  0.9817367
## 150-125  0.22 -0.09266107  0.53266107  0.2185356
## 175-125  0.20 -0.12957371  0.52957371  0.3299767
## 175-150 -0.02 -0.33266107  0.29266107  0.9976083
```

All adjusted p-values (“p adj”) for the pairwise comparisons are greater than 0.05. At the 0.05 significance level, none of the pairwise differences between firing temperatures are statistically significant.

Context: The mean density of the bricks does not differ significantly between any pair of firing temperatures. In other words, based on this experiment, the difference in changing firing temperature does not lead to a significant difference in brick density between any of the temperature groups.

## 2c

Analyze the residuals to check each of the conditions discussed in Lecture #7, with comments as appropriate. That is, check for each of the following:

- Normality – Plot a histogram of all of the residuals – Make a QQ Plot – Run the Shapiro-Wilk test
- Structure to the data – Plot the residuals vs. time
- Equal variances – Plot the residuals vs. fitted values

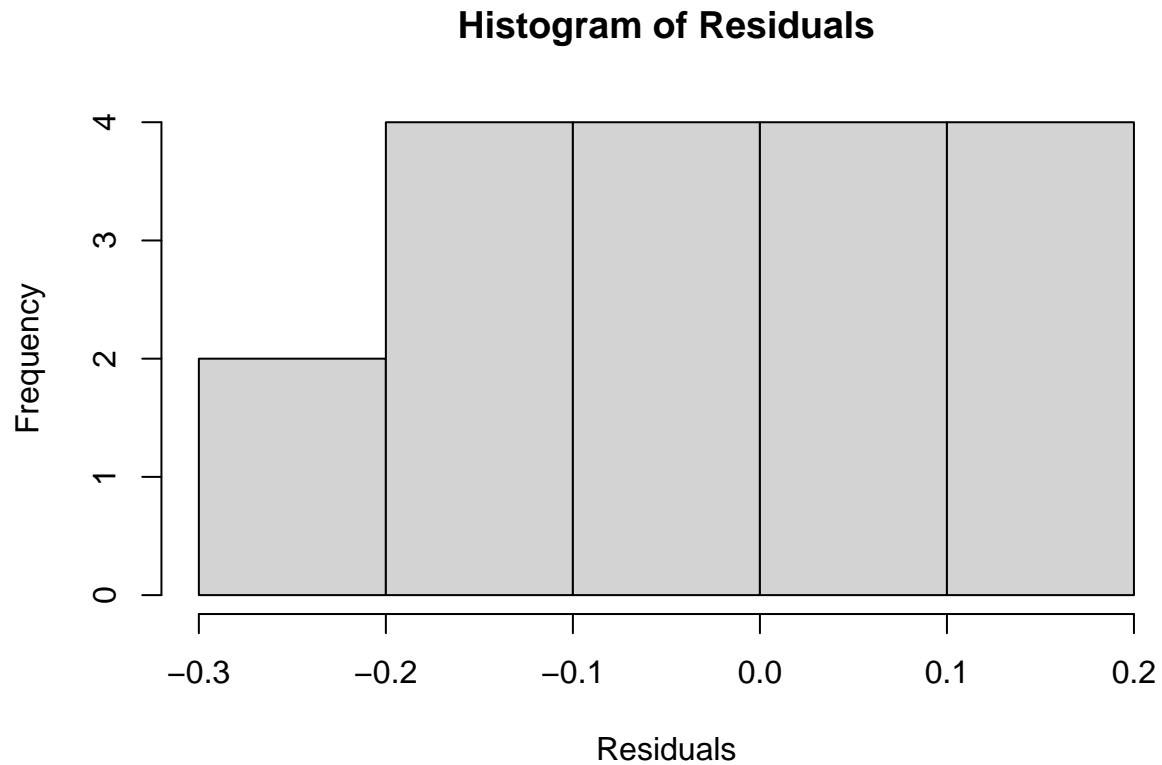
For structure to the data, we are not given the actual order of data collection; assume they were collected in the order presented in the table (from left to right, then top to bottom, i.e. in the order that you likely input them into R)

- Plot a histogram of all of the residuals

```
# calculate residuals (observation - group mean)
dens_100_resid <- dens_100 - mean(dens_100)
dens_125_resid <- dens_125 - mean(dens_125)
dens_150_resid <- dens_150 - mean(dens_150)
dens_175_resid <- dens_175 - mean(dens_175)

#Combine all residuals
dens_total_resid <- c(dens_100_resid,
```

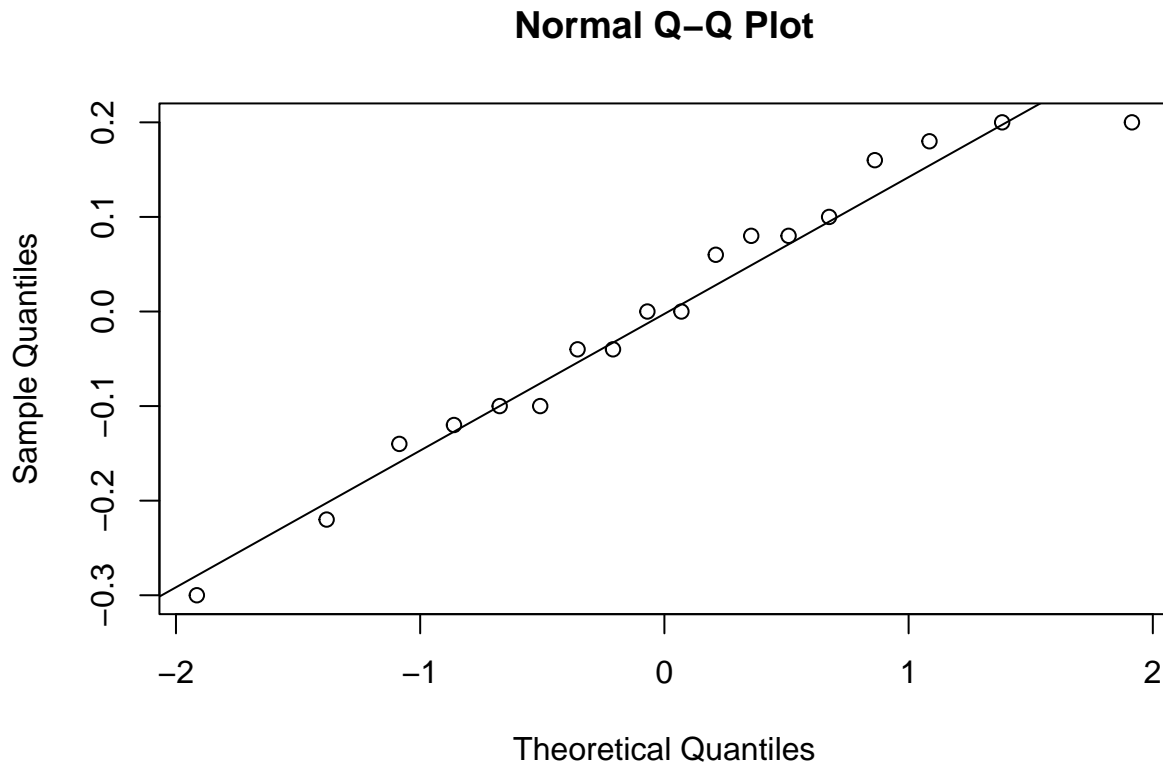
```
dens_125_resid,  
dens_150_resid,  
dens_175_resid)  
hist(dens_total_resid, main = "Histogram of Residuals", xlab = "Residuals")
```



This histogram of residuals appears roughly symmetric and centered around zero, with no extreme outliers. This suggests that the residuals are probably normally distributed.

-Make a QQ Plot

```
# using the normal probability plot  
qqnorm(dens_total_resid)  
qqline(dens_total_resid)
```



The normal Q-Q plots shows that the residuals closely follow the reference line, with only small deviations at the ends. This means that the residuals are normally distributed and the normality assumption for ANOVA is reasonable.

- Run the Shapiro-Wilk test

```
shapiro.test(mt_total_resid)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  mt_total_resid
## W = 0.97046, p-value = 0.846
```

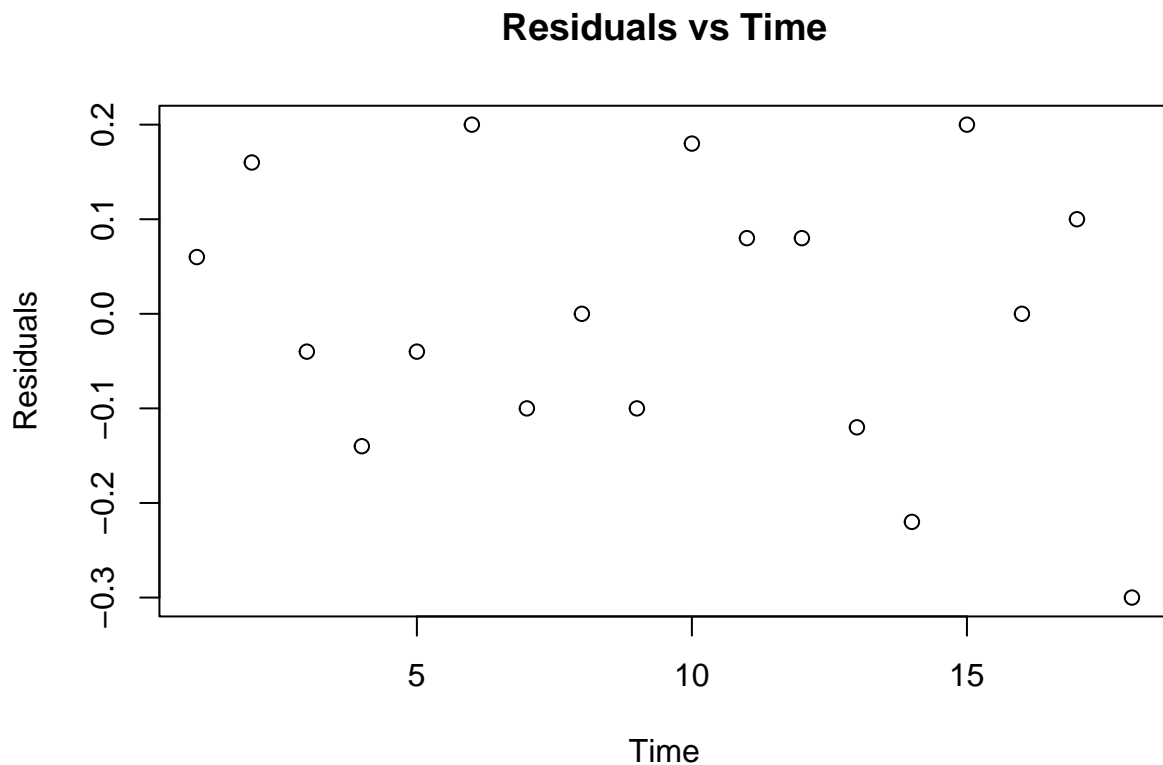
The Shapiro-Wilk normality test gives a p-value of 0.846 which is much greater than 0.05. This means there is no evidence to suggest the residuals are not normally distributed so the normality assumption for ANOVA is satisfied.

Structure to the data: - Plot the residuals vs. time

```
# fitted (predicted) values and residuals
x <- 1:length(model2$residuals)
residuals_values2 <- model2$residuals

# Plot residuals vs time
plot(residuals_values2 ~ x,
```

```
xlab = "Time",
ylab = "Residuals",
main = "Residuals vs Time")
```



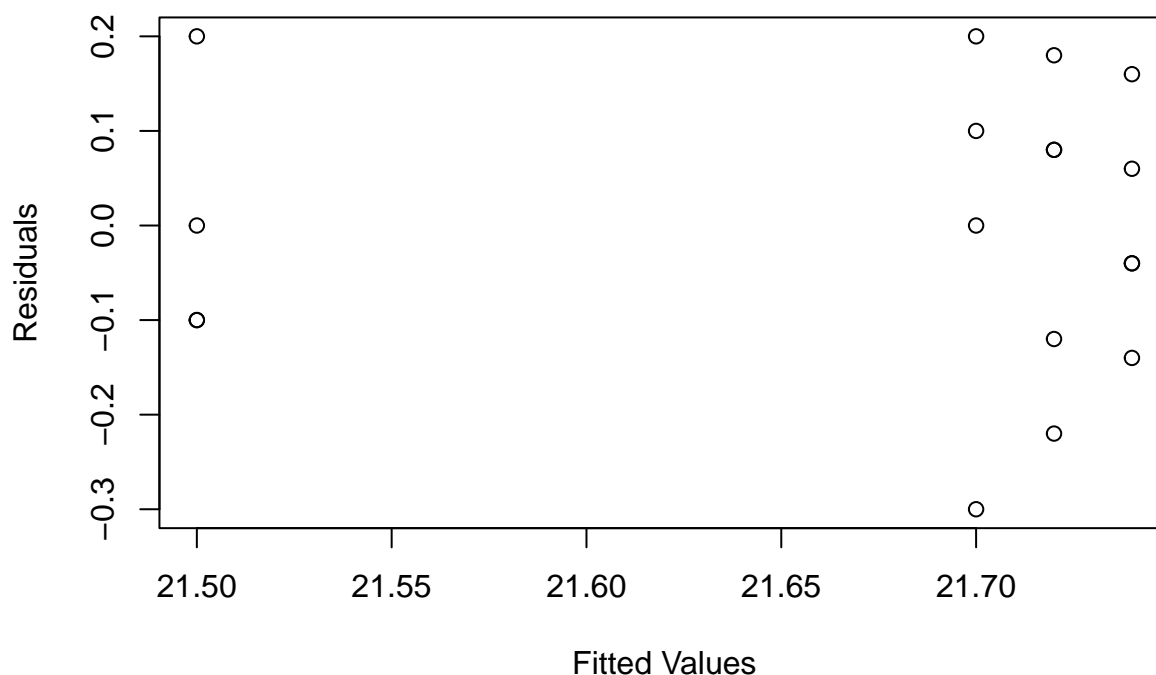
The plot of residuals versus time shows that the points are randomly scattered around zero with no obvious pattern. This means that the residuals are independent and there is no evidence of any systematic structure over time.

Equal variances - Plot the residuals vs. fitted values

```
# Extract fitted values and residuals from your ANOVA model
fitted_values2 <- model2$fitted.values
residuals_values2 <- model2$residuals

# Plot residuals vs fitted values
plot(fitted_values2, residuals_values2,
     xlab = "Fitted Values",
     ylab = "Residuals",
     main = "Residuals vs Fitted Values")
```

## Residuals vs Fitted Values



2d

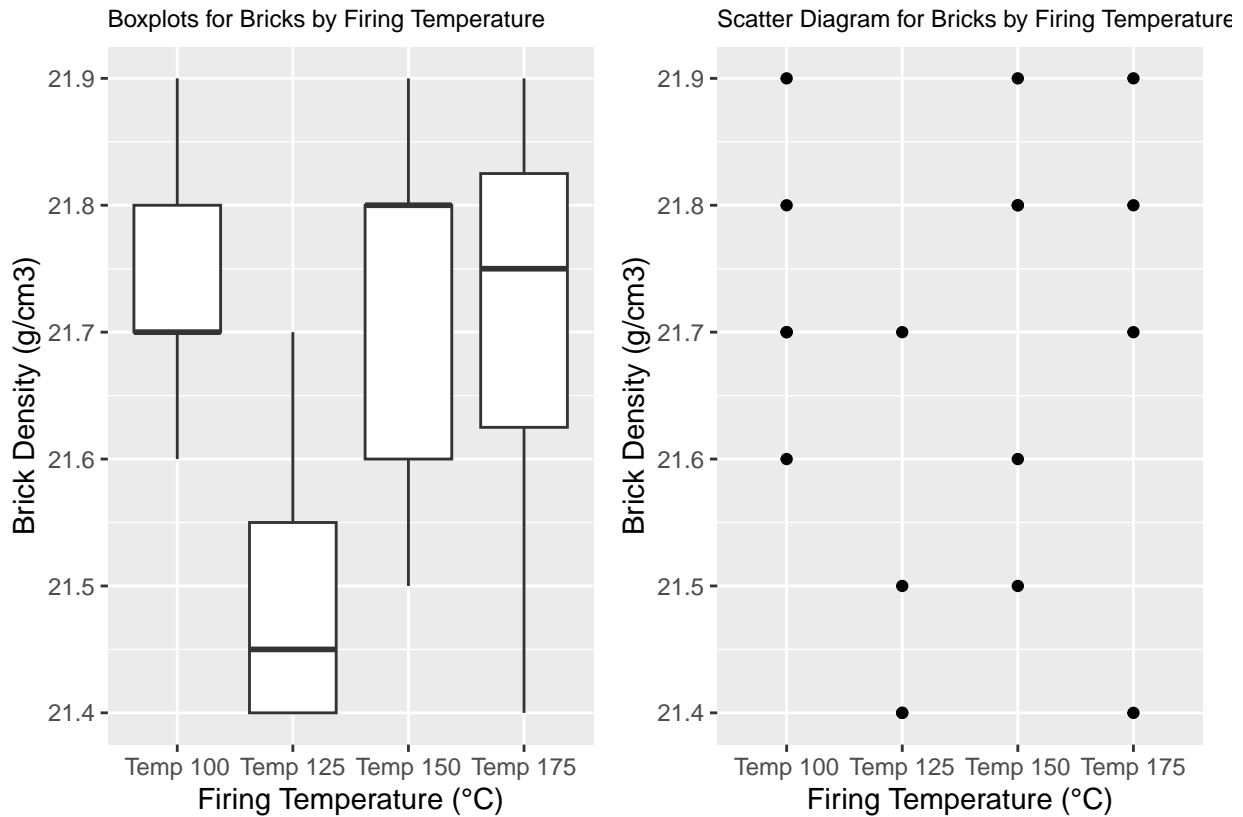
Instead making the two types of graphs we discussed in Lecture #5 for these type of data (side-by-side boxplots and a scatter diagram). Briefly describe what you observe.

```
#Boxplot
p1 <- ggplot() +
  geom_boxplot(aes(x = "Temp 100", y = dens_100)) +
  geom_boxplot(aes(x = "Temp 125", y = dens_125)) +
  geom_boxplot(aes(x = "Temp 150", y = dens_150)) +
  geom_boxplot(aes(x = "Temp 175", y = dens_175)) +
  scale_x_discrete(limits=c("Temp 100", "Temp 125", "Temp 150", "Temp 175")) +
  ggtitle("Boxplots for Bricks by Firing Temperature") +
  ylab("Brick Density (g/cm3)") +
  xlab("Firing Temperature (°C)") +
  theme(plot.title = element_text(size = 8.5)) #make the title smaller

#Scatter plot
p2 <- ggplot() +
  geom_point(aes(x = "Temp 100", y = dens_100)) +
  geom_point(aes(x = "Temp 125", y = dens_125)) +
  geom_point(aes(x = "Temp 150", y = dens_150)) +
  geom_point(aes(x = "Temp 175", y = dens_175)) +
  scale_x_discrete(limits=c("Temp 100", "Temp 125", "Temp 150", "Temp 175")) +
  ggtitle("Scatter Diagram for Bricks by Firing Temperature") +
```

```
ylab("Brick Density (g/cm3)") +
xlab("Firing Temperature (°C)") +
theme(plot.title = element_text(size = 8.5)) #make the title smaller
```

p1 + p2



Both the boxplots and scatter plot show that brick density is very similar across all firing temperatures. The medians and spreads overlap considerably, meaning it does not have strong effect on brick density.

### 3a

Do you expect that a permutation test will give you a similar p-value to that obtained by ANOVA above? Why or why not? Answer: Yes, I expect the permutation test gives me the similar p-value to the ANOVA. That's because our data seem to meet the main assumptions for ANOVA—things like normality and equal variances across groups. Since both tests are basically checking if the group means are different, and the data fits the ANOVA requirements, we should get nearly the same result from the permutation test, even though it doesn't rely on those assumptions.

### 3b

Now write code to perform a permutation test on these data, report your p-value, and comment on how it compares to the p-value from the ANOVA from the previous question.

```

set.seed(123)
F_obs <- summary(aov(density ~ temp, data = df_t))[[1]][1,4]

#Large number for the scale which gives more precise p-value
n_perm <- 1000
F_perm <- numeric(n_perm)

# Permutation test: randomly shuffle group labels and recalculate F each time
for (i in 1:n_perm) {
  permuted_temp <- sample(df_t$temp)
  perm_model <- aov(density ~ permuted_temp, data = df_t)
  F_perm[i] <- summary(perm_model)[[1]][1,4]
}

# Calculate permutation p-value
perm_pval <- mean(F_perm >= F_obs)
perm_pval

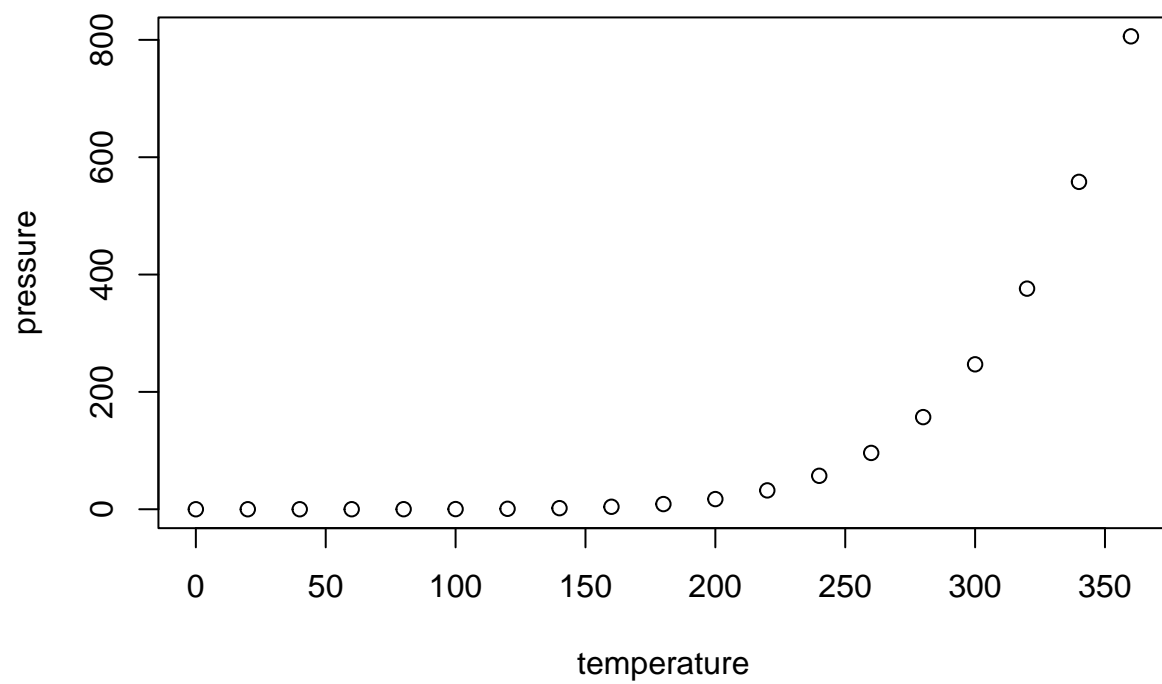
## [1] 0.165

```

The permutation test p-value 0.165. I used 1000 as my sample number which makes give us a more precise p-value. This p-value is close to the ANOVA p-value (0.157) in exercise 3.18. Based on this result, we can say that there is no statistically significant difference in brick density across the different temperature groups. Both data meet the assumptions for ANOVA.

## Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.