

Design of experiments: Lab 2

Exercise 1. [From Montgomery (2009).] During cooking, doughnuts absorb fat in various amounts. A scientist wished to learn if the amount absorbed depends on the type of fat used. For each of four fats, six batches of doughnuts were prepared. The data in the table below are the grams of fat absorbed per batch, coded by deducting 100g to give simpler figures.

T1	T2	T3	T4
64	78	75	55
72	91	93	66
68	97	78	49
77	82	71	64
56	85	63	70
95	77	76	68

- a) Read the data into R and plot it. By looking at your plot, do you think that there will be significant differences between the types of fat?

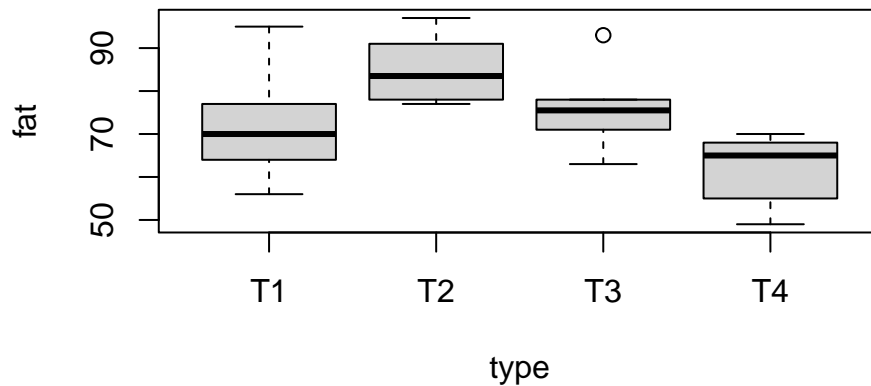
Solution:

```
library(tidyverse)
mat = matrix(c(64, 78, 75, 55,
               72, 91, 93, 66,
               68, 97, 78, 49,
               77, 82, 71, 64,
               56, 85, 63, 70,
               95, 77, 76, 68), byrow = T, ncol = 4)
colnames(mat) = c("T1", "T2", "T3", "T4")
mat = as.data.frame(mat)
df = mat %>% pivot_longer(cols = c(T1, T2, T3, T4),
                          names_to = "type",
                          values_to = "fat")
df$type = factor(df$type)
```

- b) Check that the assumptions of the one-way ANOVA model are satisfied.

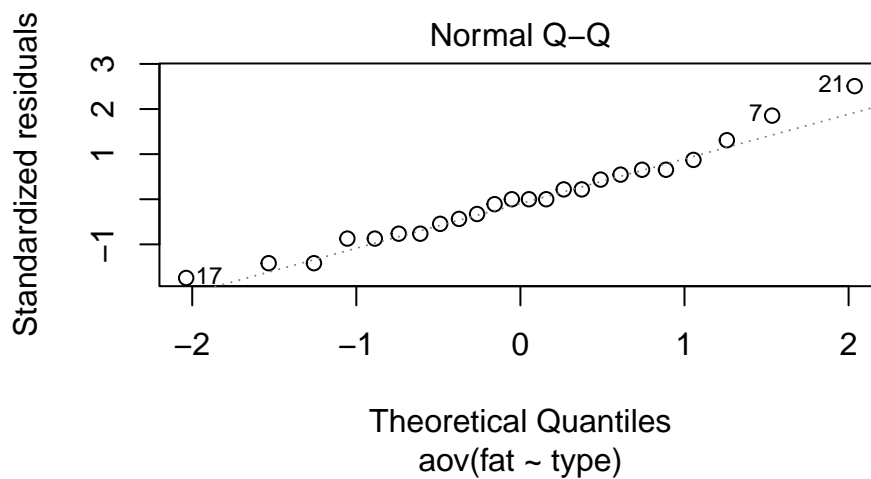
Solution: The variance of the groups doesn't look too different:

```
boxplot(fat ~ type, data = df)
```



Most points in the *qq*-plot are near the dashed line, so it seems that the residuals are approximately normal:

```
mod = aov(fat ~ type, data = df)
plot(mod, which = 2)
```



We don't have an easy way to check the assumption of independence. We can proceed with the analysis.

- c) Fit a one-way ANOVA model with the sum-to-zero constraint. Find point estimates and confidence intervals for the grand mean and the treatment effects. Comment on the results.

Solution: It seems like there will be differences between treatments. Type 2 has a big positive effect, whereas Type 4 has a strong negative effect.

```
options(contrasts = c("contr.sum", "contr.poly"))
mod = aov(fat ~ type, data = df)
dummy.coef(mod)
```

```
## Full coefficients are
##
## (Intercept):    73.75
## type:           T1    T2    T3    T4
##                -1.75  11.25  2.25 -11.75
```

```
confint(mod)
```

```
##              2.5 %    97.5 %
## (Intercept) 69.472927 78.027073
## type1      -9.158108  5.658108
## type2       3.841892 18.658108
## type3      -5.158108  9.658108
```

d) Is there evidence to claim that there are differences between types of fat at the $\alpha = 0.01$ significance level?

Solution: Yes, there is; the p -value of the global test is $< \alpha$.

```
summary(mod)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## type           3   1636    545.5     5.406 0.00688 **
## Residuals     20   2018    100.9
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

e) Run TukeyHSD to perform pairwise comparisons and comment on the results.

Solution: The only p -value that is small is that comparing type 2 to type 4.

```
TukeyHSD(mod)
```

```
##    Tukey multiple comparisons of means
##      95% family-wise confidence level
##
## Fit: aov(formula = fat ~ type, data = df)
##
## $type
##      diff      lwr      upr      p adj
## T2-T1    13 -3.232221 29.232221 0.1461929
## T3-T1     4 -12.232221 20.232221 0.8998057
## T4-T1    -10 -26.232221  6.232221 0.3378150
## T3-T2     -9 -25.232221  7.232221 0.4270717
## T4-T2    -23 -39.232221 -6.767779 0.0039064
## T4-T3    -14 -30.232221  2.232221 0.1065573
```

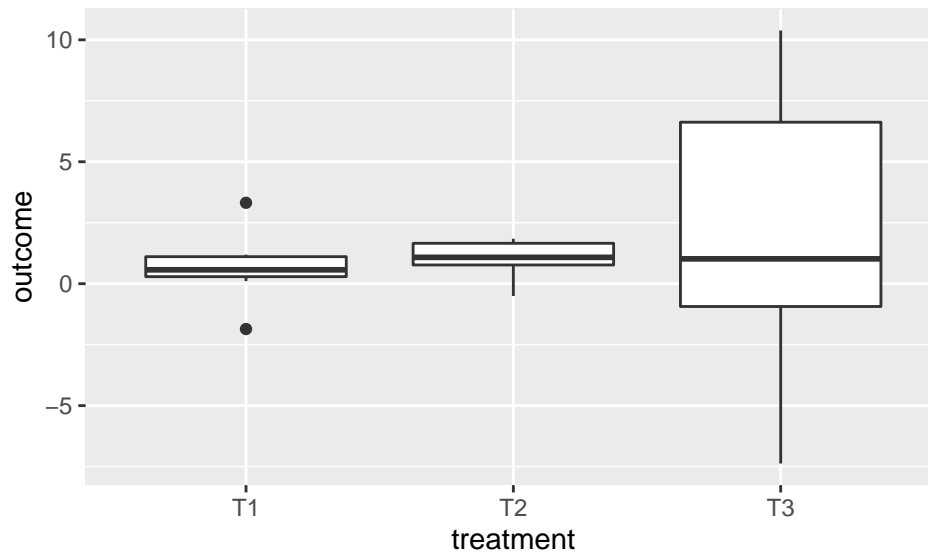
Exercise 2. In this exercise, we will work with three datasets where the assumptions of the one-way ANOVA model are not satisfied. Here is the code that reads the datasets:

```
dat1 = read.csv("http://vicpena.github.io/doe/dat1.csv")
dat2 = read.csv("http://vicpena.github.io/doe/dat2.csv")
dat3 = read.csv("http://vicpena.github.io/doe/dat3.csv")
```

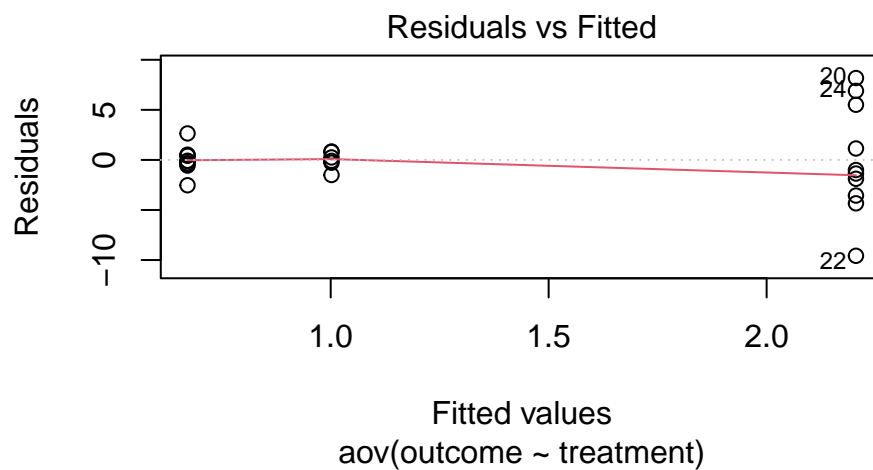
In `dat3`, the experiments were performed sequentially in an order given by the variable `order`. Identify the assumptions that are not satisfied in each case.

Solution: In the case of `dat1`, the assumption of equality of variances is violated. We can see that in a boxplot of `treatment` vs `outcome` or a plot of fitted values against observed values.

```
ggplot(dat1) +
  aes(x = treatment, y = outcome) +
  geom_boxplot()
```

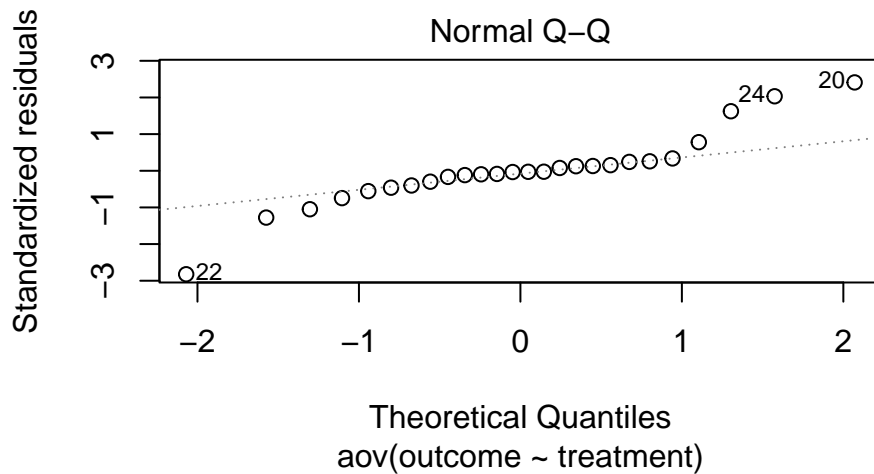


```
mod1 = aov(outcome ~ treatment, data = dat1)
plot(mod1, which = 1)
```



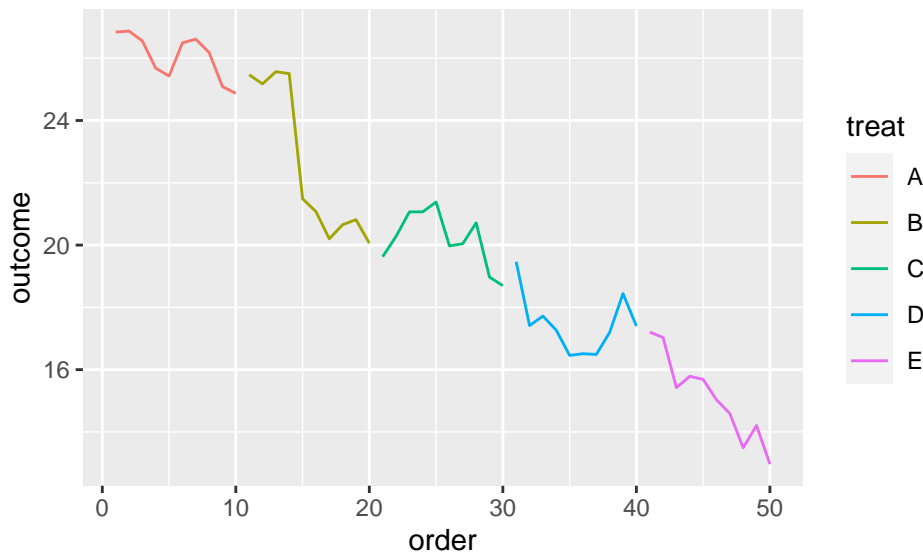
In the case of `dat2`, the assumption of normality is violated. We can especially see that in the *qq*-plot of the residuals.

```
mod2 = aov(outcome ~ treatment, data = dat2)
plot(mod1, which = 2)
```



In the case of `dat3`, the assumption of independence is violated. There is a clear order-dependence in the observations:

```
ggplot(dat3) +
  aes(x = order, y = outcome, color = treat) +
  geom_line()
```



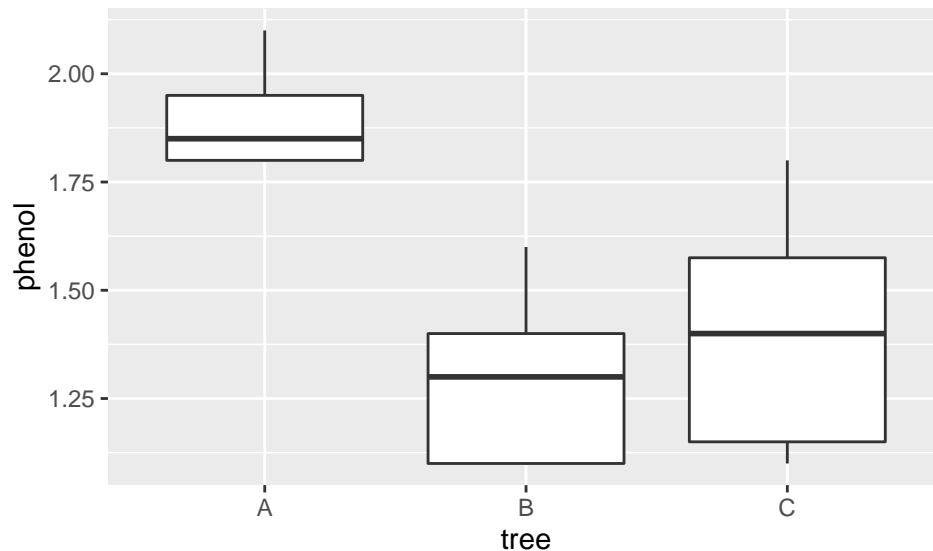
Exercise 3. In the process of turning wood into paper sheets, phenol waste is produced. A group of technicians working at a paper company want to compare the average phenol waste produced by 3 different species of trees. The code below reads in the data, which contains the percentage of phenol waste produced in 15 experiments.

```
phenol = read.csv("http://vicpena.github.io/doe/phenol.csv")
```

a) Are there differences between tree species?

Solution: There are significant differences between trees at the $\alpha = 0.05$ significance level.

```
ggplot(phenol) +
  aes(x = tree, y = phenol) +
  geom_boxplot()
```



```
mod = aov(phenol ~ tree, data = phenol)
summary(mod)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## tree           2   0.90  0.4500    8.437 0.00515 **
## Residuals     12   0.64  0.0533
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We can run a TukeyHSD to see where the differences are:

```
TukeyHSD(mod)
```

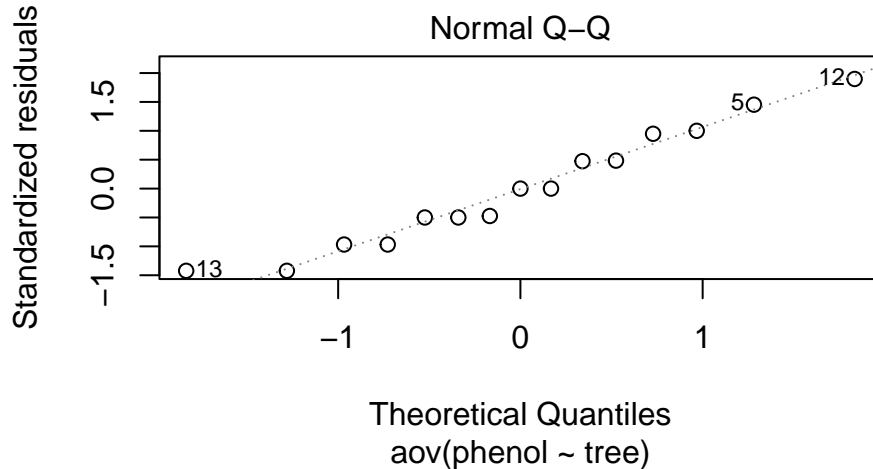
```
##    Tukey multiple comparisons of means
##      95% family-wise confidence level
##
## Fit: aov(formula = phenol ~ tree, data = phenol)
##
## $tree
##      diff      lwr      upr      p adj
## B-A -0.6 -1.0133037 -0.1866963 0.0058028
## C-A -0.5 -0.8977016 -0.1022984 0.0146816
## C-B  0.1 -0.2730772  0.4730772 0.7594319
```

There are significant differences between A and B and A and C at $\alpha = 0.05$.

- b) What assumptions did you make to draw your conclusions in part a)? Are they reasonable in this context?

Solution: As usual, we make assumptions of normality, independence, and equality of variances. The latter seems to be satisfied by looking at the boxplot in part a). The qq-plot looks good, so the assumption of normality doesn't seem to be violated. We don't have enough information to check independence.

```
plot(mod, which = 2)
```



- c) The technicians ran all the experiments for type A first, then those for type B, and then those for type C. Do you think that's reasonable, or would you design the experiment differently?

Solution: It's not ideal. It would be best to run the experiments in random order to avoid temporal dependence.

- d) Actually, the design was originally balanced, but the technicians excluded three observations from the analysis (with phenol waste percentages 3.6, 4.1, and 3.5). Do you think this is reasonable?

Solution: It depends. Phenol waste percentages of 3.6, 4.1, and 3.5 are certainly abnormal values compared to the other data. If they were excluded because they made a mistake in the experiment, it seems reasonable to exclude them. If not, the technicians should find out what's wrong with those observations.

Exercise 4. [Ott (1973)] In the process of manufacturing an electronic circuit, an engineer noticed that there was more variability than expected between circuits. Having studied the problem, the engineer concluded that one of the sources of variability could be a flat piece of insulating ceramic that is placed in the circuits.

This flat piece is obtained from ceramic sheets that are purchased from an outside supplier. Each sheet is cut into small pieces, each of which is placed in a circuit. The engineer suspects that there is variation between sheets and that this difference affects the outcome.

You can read in the data with the following command:

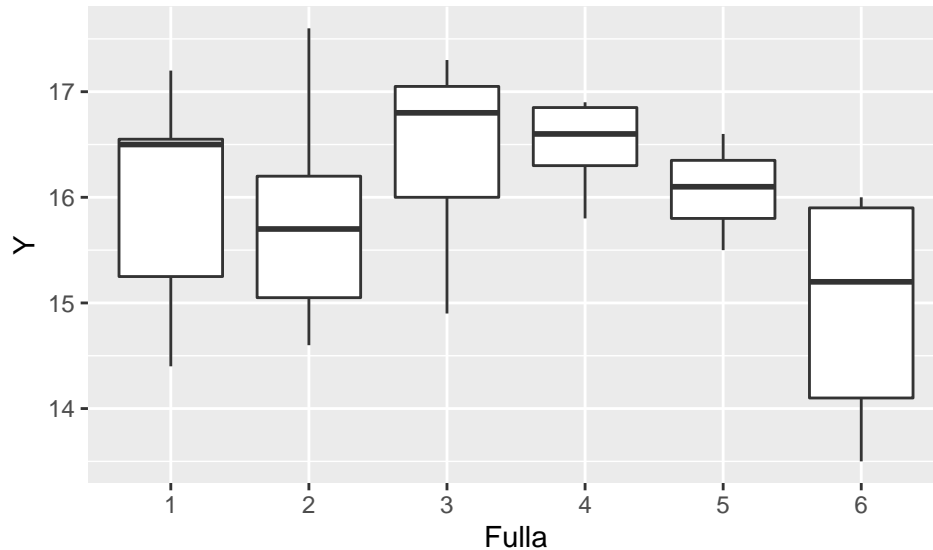
```
ott = read.csv2("http://vicpena.github.io/doe/Ott_Case_13_1.csv")
```

To analyze whether the suspicion was founded, 6 sheets were taken and cut into 7 pieces each. Then, a circuit was assembled in each of the 42 pieces.

- a) Plot the data and draw some preliminary conclusions about the influence of the sheet on the outcome.

Solution It seems that there are some differences between sheets.

```
ott$Fulla = as.factor(ott$Fulla)
ggplot(ott) +
  aes(x = Fulla, y = Y) +
  geom_boxplot()
```



b) Does the ceramic sheet influence the outcome?

```
mod = aov(Y ~ Fulla, data = ott)
summary(mod)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## Fulla      5  11.16   2.2312    3.117 0.0194 *
## Residuals 36   25.77   0.7158
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p -value is significant at $\alpha = 0.05$.

We can run a TukeyHSD to see where the differences are:

```
TukeyHSD(mod)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = Y ~ Fulla, data = ott)
##
## $Fulla
##           diff           lwr           upr           p adj
## 2-1 -0.18571429 -1.5462837  1.1748551 0.9983731
## 3-1  0.48571429 -0.8748551  1.8462837 0.8884092
## 4-1  0.55714286 -0.8034265  1.9177123 0.8183341
## 5-1  0.11428571 -1.2462837  1.4748551 0.9998467
```



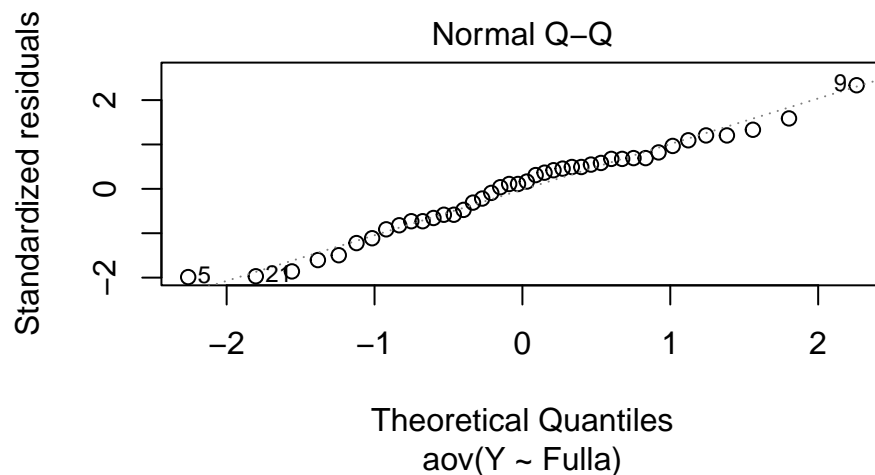
```
## 6-1 -1.00000000 -2.3605694 0.3605694 0.2577389
## 3-2 0.67142857 -0.6891408 2.0319980 0.6758457
## 4-2 0.74285714 -0.6177123 2.1034265 0.5768797
## 5-2 0.30000000 -1.0605694 1.6605694 0.9848403
## 6-2 -0.81428571 -2.1748551 0.5462837 0.4780934
## 4-3 0.07142857 -1.2891408 1.4319980 0.9999850
## 5-3 -0.37142857 -1.7319980 0.9891408 0.9616346
## 6-3 -1.48571429 -2.8462837 -0.1251449 0.0255115
## 5-4 -0.44285714 -1.8034265 0.9177123 0.9215549
## 6-4 -1.55714286 -2.9177123 -0.1965735 0.0170585
## 6-5 -1.11428571 -2.4748551 0.2462837 0.1621012
```

Most p -values aren't significant at $\alpha = 0.05$. There are significant differences between sheets 4 and 6 and sheets 3 and 6.

c) Verify that the assumptions of the model are satisfied.

Solution Looking at the boxplot in part a), the assumption of equality of variances may be a little suspect. The qq-plot looks fine, so the hypothesis of normality seems reasonable. We don't have data to verify the hypothesis of independence.

```
plot(mod, which = 2)
```



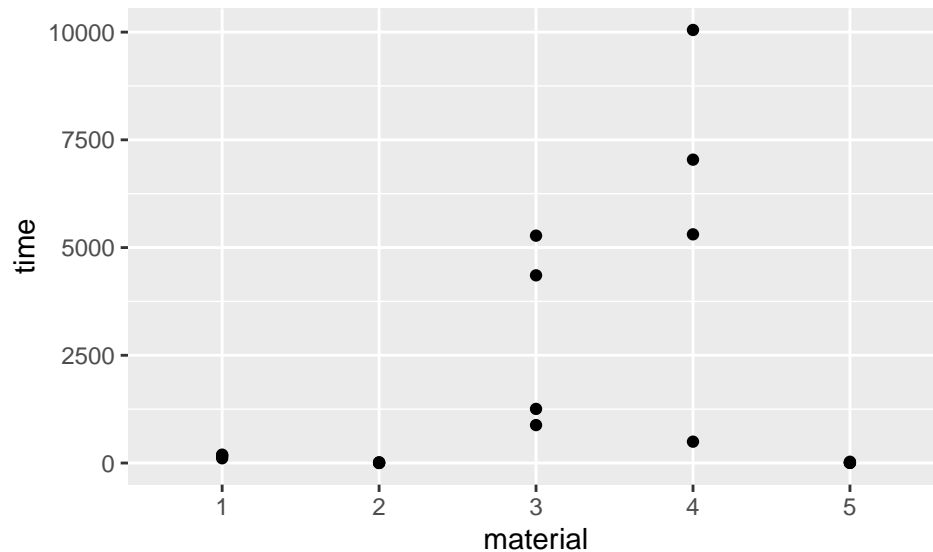
Exercise 5. [Montgomery (2009)] An experiment is conducted to investigate the effectiveness of 5 insulating materials. Four test pieces of each material are taken and subjected to a high potential difference to accelerate their time to failure. The times to failures (in minutes) are shown in the following table:

Material	Time to failure			
1	110	157	194	178
2	1	2	4	18
3	880	1256	5276	4355
4	495	7040	5307	10050
5	7	5	29	2

a) Plot the data and comment on what you see.

Solution: There seem to be differences between materials. It also seems clear that the variance depends on the material.

```
ins = matrix( c(1 , 110 , 157 , 194 , 178 ,
                2 , 1 , 2 , 4 , 18 ,
                3 , 880 , 1256 , 5276 , 4355 ,
                4 , 495 , 7040 , 5307 , 10050 ,
                5 , 7 , 5 , 29 , 2), nrow = 5, byrow = T)
colnames(ins) = c("material", "r1", "r2", "r3", "r4")
ins = as.data.frame(ins)
ins$material = as.factor(ins$material)
ins = ins %>% pivot_longer(c(r1, r2, r3, r4),
                           names_to = "replicate",
                           values_to = "time") %>%
  select(-replicate)
ggplot(ins) +
  aes(x = material, y = time) +
  geom_point()
```



b) Do the materials have the same effect on the average time to failure?

Solution: If we run `aov`, we'll find a significant p -value at $\alpha = 0.05$. However, the assumptions of the model aren't satisfied; see part c).

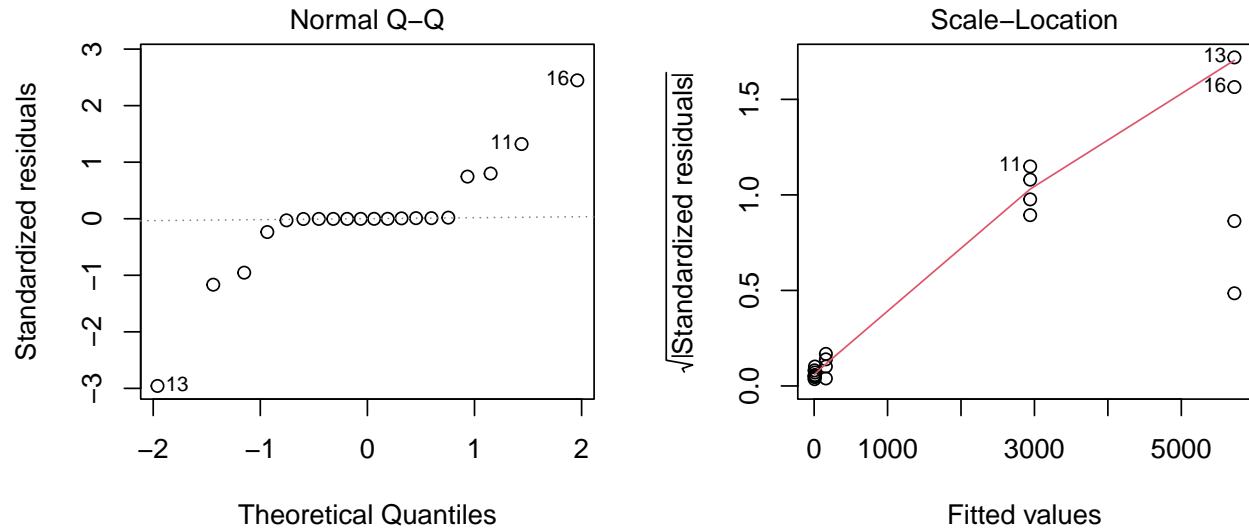
```
mod = aov(time ~ material, data = ins)
summary(mod)
```

```
##           Df    Sum Sq Mean Sq F value Pr(>F)
## material    4 103191489 25797872   6.191 0.00379 **
## Residuals   15  62505657  4167044
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

c) Check the model assumptions. Do you think they are satisfied?

Solution: The assumptions of normality and equality of variance are clearly not satisfied.

```
par(mfrow = c(1, 2))
plot(mod, which = 2)
plot(mod, which = 3)
```



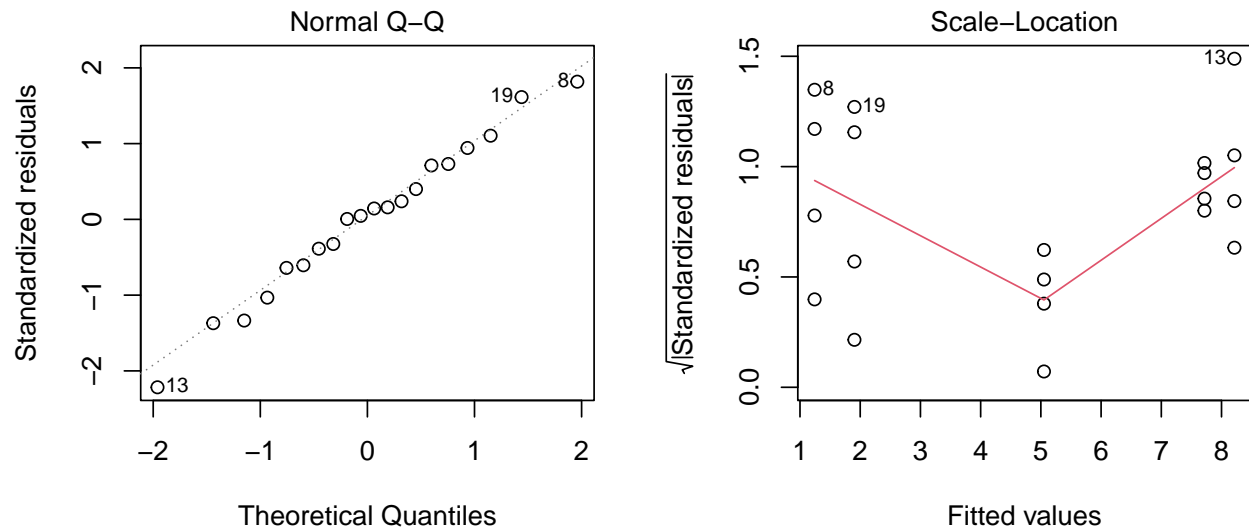
- d) Try out taking a logarithmic transformation of the time to failure and running the analysis again. Comment on the results.

Solution: The residuals look good after taking the logarithmic transformation. The F test tells us that there are differences between materials.

```
mod_log = aov(log(time) ~ material, data = ins)
summary(mod_log)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## material     4  165.06   41.26   37.66 1.18e-07 ***
## Residuals   15   16.44    1.10
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
par(mfrow = c(1, 2))
plot(mod_log, which = 2)
plot(mod_log, which = 3)
```



We can find where the differences are with TukeyHSD:

```
TukeyHSD(mod_log)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = log(time) ~ material, data = ins)
##
## $material
##      diff      lwr      upr    p adj
## 2-1 -3.8091386 -6.0948210 -1.5234562 0.0009669
## 3-1  2.6648112  0.3791288  4.9504936 0.0188866
## 4-1  3.1624158  0.8767334  5.4480982 0.0051317
## 5-1 -3.1476442 -5.4333266 -0.8619618 0.0053339
## 3-2  6.4739498  4.1882674  8.7596322 0.0000025
## 4-2  6.9715545  4.6858720  9.2572369 0.0000010
## 5-2  0.6614944 -1.6241880  2.9471768 0.8948496
## 4-3  0.4976046 -1.7880778  2.7832870 0.9594662
## 5-3 -5.8124554 -8.0981378 -3.5267730 0.0000094
## 5-4 -6.3100600 -8.5957424 -4.0243776 0.0000034
```

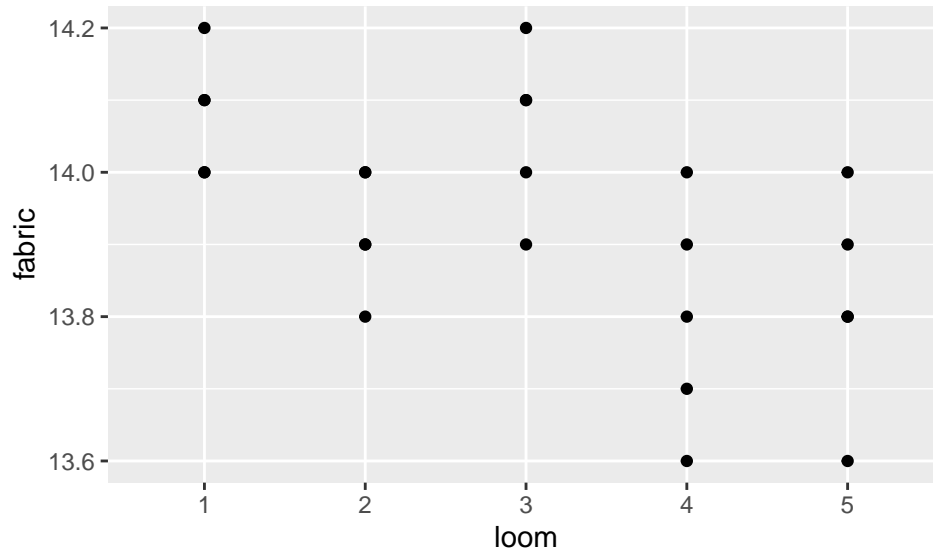
Exercise 6. A textile factory has 5 looms. Each loom is assumed to provide the same amount of fabric per minute. To investigate this hypothesis, the company recorded the Kg of fabric processed in one minute at the looms at different times. The data obtained are the following:

Loom	Output [Kg/min]				
1	14.0	14.1	14.2	14.0	14.1
2	13.9	13.8	13.9	14.0	14.0
3	14.1	14.2	14.1	14.0	13.9
4	13.6	13.8	14.0	13.9	13.7
5	13.8	13.6	13.9	13.8	14.0

a) Plot the data and fit a one-way ANOVA model. Check the assumptions of the model.

Solution: I'm not plotting the results here for concreteness, but the residuals look fine.

```
mat = matrix( c(1, 14.0, 14.1, 14.2, 14.0, 14.1,
                2, 13.9, 13.8, 13.9, 14.0, 14.0,
                3, 14.1, 14.2, 14.1, 14.0, 13.9,
                4, 13.6, 13.8, 14.0, 13.9, 13.7,
                5, 13.8, 13.6, 13.9, 13.8, 14.0),
              nrow = 5, byrow = T)
colnames(mat) = c("loom", "r1", "r2", "r3", "r4", "r5")
df = as.data.frame(mat)
df = df %>% pivot_longer(c(r1, r2, r3, r4, r5),
                        names_to = "rep",
                        values_to = "fabric") %>%
  select(-rep)
df$loom = factor(df$loom)
ggplot(df) +
  aes(x = loom, y = fabric) +
  geom_point()
```



b) Is there evidence to reject the hypothesis that the looms produce the same amount of fabric?

Solution: There are significant differences at the $\alpha = 0.05$ significance level.

```
mod_loom = aov(fabric ~ loom, data = df)
summary(mod_loom)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## loom         4  0.3416   0.0854     5.77 0.00296 **
## Residuals    20  0.2960   0.0148
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

However, one could argue that the differences may not be practically relevant. Let's take a look at TukeyHSD, which gives us intervals quantifying the estimated differences between looms (as well as p -values):

```
TukeyHSD(mod_loom)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = fabric ~ loom, data = df)
##
## $loom
##      diff      lwr      upr    p adj
## 2-1 -0.16 -0.3902379 0.070237895 0.2669972
## 3-1 -0.02 -0.2502379 0.210237895 0.9989043
## 4-1 -0.28 -0.5102379 -0.049762105 0.0125769
## 5-1 -0.26 -0.4902379 -0.029762105 0.0221607
## 3-2  0.14 -0.0902379 0.370237895 0.3904049
## 4-2 -0.12 -0.3502379 0.110237895 0.5382897
## 5-2 -0.10 -0.3302379 0.130237895 0.6944089
## 4-3 -0.26 -0.4902379 -0.029762105 0.0221607
## 5-3 -0.24 -0.4702379 -0.009762105 0.0384679
## 5-4  0.02 -0.2102379 0.250237895 0.9989043
```

The estimated differences are, at most, of 0.5 kg. We should talk to the workers at the factory and ask them if these differences are practically relevant or not.

- c) Find an estimate of the variability between looms and compare it to the variability within looms.

Solution: The variability between looms can be estimated with the mean square of `loom` in the ANOVA table: it is 0.0854. Analogously, the variability within looms can be estimated with the mean square `Residuals` from the ANOVA table, which is 0.0148.