1. A trendy wine bar set up an experiment to evaluate the quality of 3 different wines. Five fine connoisseurs of wine were asked to taste each of the wine and give it a rating between 0 and 10. The order of tasting was randomized and the judges did not know which wine they were drinking.
The following table displays the collected data:

	wine 1	wine 2	wine 3
person 1	1	7	5
person 2	0	4	0
person 3	1	6	4
person 4	1	5	2
person 5	1	8	10

We use the following model:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij},$$

where Y_{ij} are the ratings and α_i, β_j the (fixed) effects of wine type and person, respectively. We use the sum-to-zero constraint

$$\sum_{i=1}^{3} \alpha_i = \sum_{j=1}^{5} \beta_j = 0$$

and the standard assumptions for the errors ϵ_{ij}.

The following R-output is available:

```
> options(contrasts = c("contr.sum", "contr.sum"))
> fit <- aov(y ~ wine + person, data = wine_tasting)
> summary(fit)
            Df Sum Sq Mean Sq F value Pr(>F)
wine         2   69.7    34.9   10.90 0.0052 **
person       4   42.0    10.5    3.28 0.0717 .
Residuals    8   25.6     3.2
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> coef(fit)
(Intercept)       wine1       wine2     person1     person2
    3.7e+00    -2.9e+00     2.3e+00     6.7e-01    -2.3e+00
    person3     person4
   -5.6e-17    -1.0e+00
> dummy.coef(fit)
Full coefficients are

(Intercept):          3.7
wine:               wine 1    wine 2    wine 3
                     -2.87      2.33      0.53
person:                  1         2         3         4
                   6.7e-01 -2.3e+00 -5.6e-17 -1.0e+00

(Intercept):
wine:

person:                  5
                   2.7e+00
```

a) What design do we have here? What is the role of the different factors?

b) Does wine type have an effect on rating? Use the global test. State the null hypothesis with respect to the corresponding parameters, the p-value and the test result.

c) What is the estimated rating difference between wine 1 and wine 2?

d) Should we really include the effect of the person (β_j) in the model or not? Motivate your answer.

e) Your colleague wants to run the following code in R

```
> fit2 <- aov(y ~ wine * person, data = wine_tasting)
> summary(fit2)
```

What model is he using? Will he be able to perform statistical tests?

f) If we assume that raters were randomly selected, we could model them as random effects. Determine a 95% confidence interval for the standard deviation of this random effect using the outputs below.

```
> library(lmerTest)
> fit3 <- lmer(y ~ wine + (1 | person), data = wine_tasting)
> fit4 <- lmer(y ~ person + (1 | wine), data = wine_tasting)
> confint(fit3, oldNames = FALSE)

                      2.5 % 97.5 %
sd_(Intercept)|person   0.0    3.5
sigma                   1.1    2.7
(Intercept)             1.9    5.5
wine1                  -4.1   -1.6
wine2                   1.7    3.2

> confint(fit4, oldNames = FALSE)

                    2.5 % 97.5 %
sd_(Intercept)|wine  0.91   6.33
sigma                1.03   2.32
(Intercept)          0.19   7.14
person1             -0.94   2.27
person2             -3.94  -0.73
person3             -1.60   1.60
person4             -2.60   0.60

> rand(fit3)

Analysis of Random effects Table:
       Chi.sq Chi.DF p.value
person   2.03      1     0.2

> rand(fit4)

Analysis of Random effects Table:
     Chi.sq Chi.DF p.value
wine   6.14      1    0.01 *
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

2. A pharmaceutical company wants to study the effect of alcohol consumption in conjunction with two types of drugs, A and B. They consider the following treatments:

1. Control
2. Drug A alone
3. Drug A and alcohol consumption 1 hour before
4. Drug A and alcohol consumption 1 hour after
5. Drug B alone
6. Drug B and alcohol consumption 1 hour before
7. Drug B and alcohol consumption 1 hour after

A completely randomized design was used. Every individual in the experiment was given to test one treatment only and the effect of the drug was measured on some scale and recorded as variable Y.

a) We want to ask precise questions about the data and use contrasts to do so. Propose contrasts to test the following questions:
- $L1$: The difference between the drug A and the drug B
- $L2$: The effect of alcohol on drug A
- $L3$: The difference between taking alcohol before and after

b) Are the previous 3 contrasts orthogonal to each other? Justify your answer.

c) What do the two following contrasts test? Explain with words.
- $L4$: $(6, -1, -1, -1, -1, -1, -1)$
- $L5$: $(0, 0, +1, -1, 0, -1, +1)$

d) We fit the following model in R

```
> fit <- aov(y ~ treatment, data = drug)
> summary(fit)
             Df  Sum Sq Mean Sq F value Pr(>F)
treatment     6 4738965  789828   238.5 <2e-16 ***
Residuals   833 2758944    3312
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

How many people took part in this study?

e) What procedure would you suggest if someone asks you to perform all possible pairwise comparisons between the different treatments?

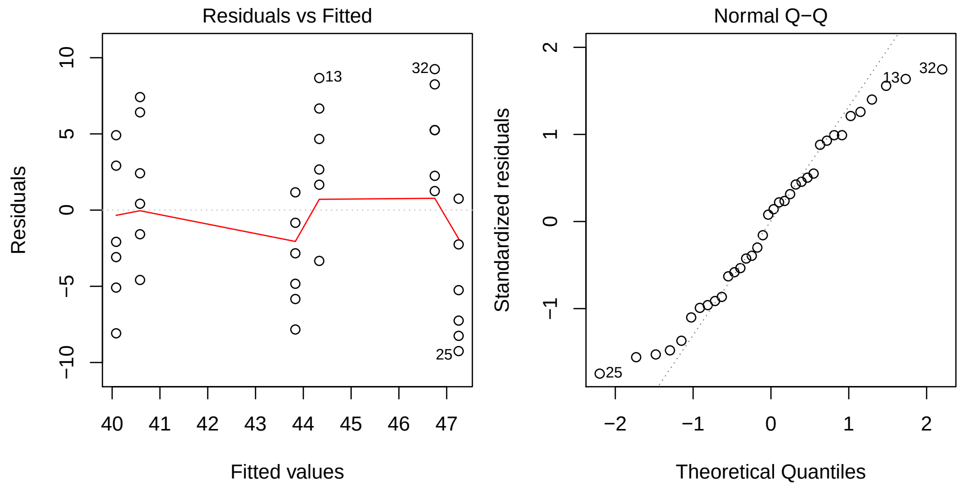
f) Say you want to test all the 5 contrasts from above. Is it necessary to adjust the corresponding p-values and if yes, how?

3. A pizzeria wants to optimize its least sold Pizza Margherita to guarantee a maximum taste experience. To find the best combination of baking temperature and baking time, they perform an experiment with two factors (temp = 180 °C, 210 °C and 240 °C, time = 10 min and 15 min). For every combination, 6 pizzas get baken and judged on a scale of 1 to 100 where 100 corresponds to maximum taste experience. As analysis a two-way anova is performed with the following output:

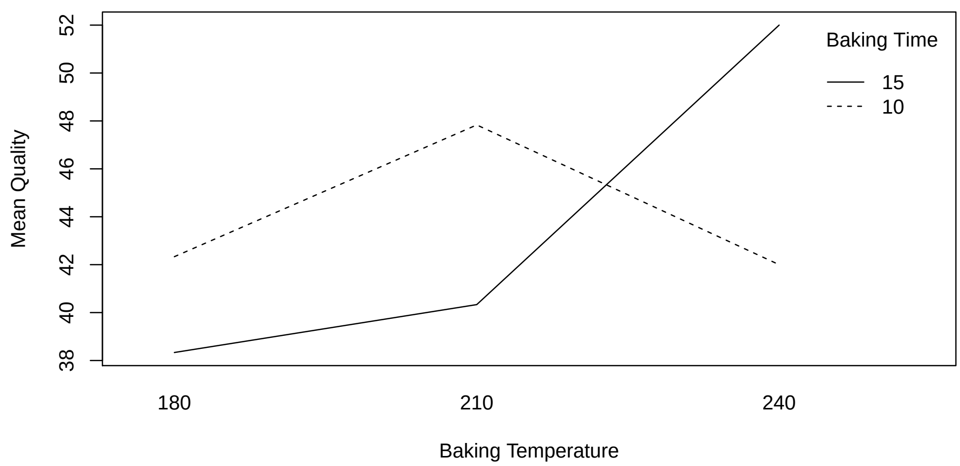
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
temp	2	268	134.0	4.25	0.023
time	1	2	2.2	0.07	0.791
Residuals	32	1009	31.5		

a) From the R-Output above, what is the estimated error variance?

b) What are the standard assumptions about the errors? Are these assumptions fullfilled? Motivate your answer with the help of the following residual plots.



c) Have a look at the following plot. What does the plot show? Explain why the model above can lead to wrong conclusion?



d) Have a look at the following output. Complete the missing parts in the first 3 rows.

```
             Df  Sum Sq  Mean Sq  F value   Pr(>F)
temp          2                     8.13   0.0015
time          1       2     2.25           0.7145
temp:time           514            15.60  2.3e-05
Residuals    30     495    16.49
```

e) From the output above, which is the final model you would choose and why? Write down your selected model in the form $Y_{ij} = \mu + \alpha_i + \cdots$.

f) Assume you want to repeat the analysis after a year. However, some data got lost and the number of judgements for each pizza is therefore not the same anymore. Does the loss of data have an influence on the calculation of the ANOVA table? What R-function would you use? Motivate your answers.

4. Misc

1) You are given the following experimental design with two block factors X and Y each having 4 levels and a treatment factor A with 4 levels: A_1, A_2, A_3, A_4. Which type of design is this?

		Blocks X			
		1	2	3	4
Blocks Y	1	A_3	A_2	A_4	A_1
	2	A_1	A_3	A_2	A_4
	3	A_4	A_1	A_3	A_2
	4	A_2	A_4	A_1	A_3

   a) Split-plot design
   b) 2^3 design
   c) Latin square design
   d) Balanced incomplete block design
   e) Completely randomized design
   f) None of the previous designs

2) A toothpaste company is testing 10 new toothpaste types. 15 participants (=blocks) have been selected. You wish to run BIBD with block size 3 such that each toothpaste type is being tested a total of 6 times. Is this possbible?

   a) Yes
   b) No
   c) Not enough information to make a statement

3) The toothpaste factory has selected 4 out of 10 types from the previous test. They are now considering these 4 types of toothpaste and 3 types of packaging. 60 participants have been selected for the experiment, and each participant is supposed to test and rate every packaging of exactly one toothpaste type on a 1-5 scale. Which type of design is this?

   a) Split-plot design with participants as whole-plots
   b) Split-plot design with toothpaste as whole-plots
   c) Split-plot design with packagings as whole-plots
   d) None of the previous designs

4) Consider a (balanced) one-way ANOVA model. What happens to the 95%-quantile of the F-distribution of the global test if we increase the number of observations (but keep the design fixed otherwise)?

   a) The quantile gets larger
   b) The quantile gets smaller
   c) The quantile stays the same
   d) No statement possible

5) We have an unbalanced data-set with two factors A and B and fit the model `aov(y ~ A * B)` in R.

   a) Type II and type III sum of squares coincide for factor B.
   b) Type I and type II sum of squares coincide for factor A.

c) The sum of squares of the interaction are the same for all types (I - III).

6) The following table contains *mean values* of an experiment with two crossed factors A and B

Level	B_1	B_2	B_3
A_1	4	8	6
A_2	10		12

What value would you put in the missing cell if you assume an additive model (that is, no interaction between A and B)?

a) 11

b) 13

c) 14

d) 16