Exercise 29

In an experiment to analyze the quality of insecticides the same number of insects were exposed (multiple times) to the sprays A, B, C, D, E, and F, respectively. The following table shows the absolute frequencies of insects surviving the sprays (counts are already ordered by size):

Spray	Count
A	7 10 10 12 13 14 14 14 17 20 20 23
B	7 11 11 13 14 16 17 17 17 19 21 21
C	0 0 1 1 1 1 2 2 3 3 4 7
D	2 3 3 4 4 5 5 5 5 5 6 12
E	1 1 2 3 3 3 3 4 5 5 6 6
F	9 10 11 13 13 15 15 16 22 24 26 26

(a) Visualize the data by an appropriate plot, e.g. with plotting functions from the package `ggplot2`. What conclusions can you draw about the effectiveness of the insecticides?

(b) Perform a one-way ANOVA (at significance level $\alpha = 0.01$) to investigate whether there is a difference in the effects of the insecticides.

(c) Which of the ANOVA model assumptions might be violated here? Explain your conclusions, also in conjunction with your plot from (a). What about using the transformation $\log(1+\ )$ on the data? Would it be helpful in this regard?

Exercise 30

Consider the data set `shafii.rapeseed` in the package `agridat`.

(a) For now consider only the factor variables Location and Genotype. Perform a two-way ANOVA to investigate their effect on the (response) variable yield. Complement your analysis by additionally visualizing the data in an appropriate way (e.g. with functions from the package `ggplot2`).

(b) Now additionally include the factor variable Year in your model and investigate the effect of all 3 factors, and their interactions (if necessary) on the variable yield. Again, visualize the data in an appropriate way.

Exercise 31

An field trial is conducted to investigate whether yields of turnips (`yield`) grown for winter fodder depend on the factor variables genotype (`gen`) and planting density (`density`). The trial is set up as a balanced design. The factor variable `gen` has levels `Barkant` and `Marco`, the factor variable `density` has levels `1`, `2`, `4` and `8`, measured in kg/ha.

turnip field trial: yield vs planting density, given genotype
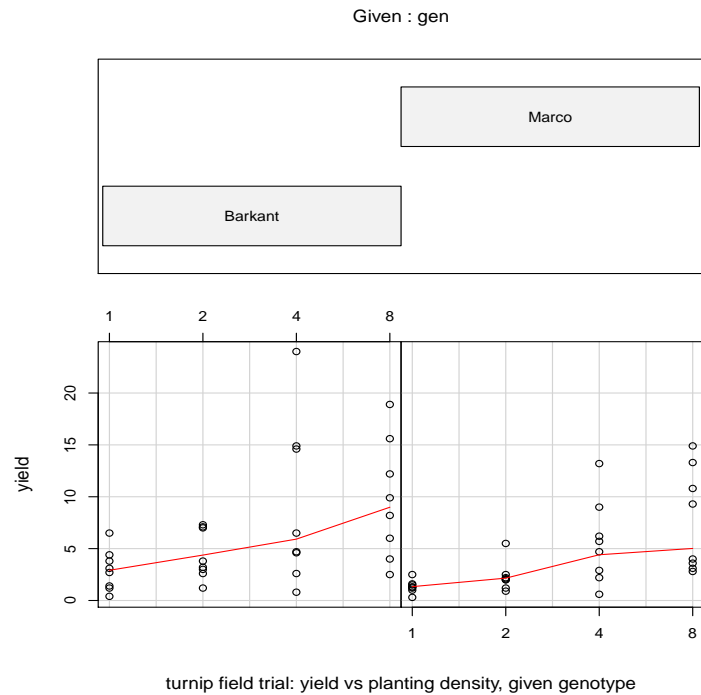
A two-way ANOVA (at significance level $\alpha = 0.05$) is conducted to analyze whether `gen` and `density` have a significant effect on `yield`. The following (incomplete) output was obtained.

```
Analysis of Variance Table

Response: yield
            Df  Sum Sq Mean Sq F value     Pr(>F)
density      3  470.38 156.793  8.4565 0.0001002 ***
gen          1   83.95  83.951    (??) 0.0377628 *
density:gen  3    8.65    (?)           0.1555 0.9257472
Residuals   56 1038.30  18.541
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(a) Write down the corresponding model in terms of a mathematical formula.

(b) Compute the missing quantities (?) and (??) in the R output. On how many turnips is the analysis based?

(c) Interpret the result of the ANOVA. What conclusions can you draw about the effect of `gen` and `density`, and in particular about the effect of the term `density:gen`, on the turnip yield? For your interpretation also take into account above graphical representation of the data.