EXAM COVER SHEET – Version B, Problem 1

Module:	**D2, Design and Analysis of Experiments**
Date of exam:	19.01.2021, 1.00 – 3.00pm
Duration:	2x 45 min
Type of exam:	Open book: Distributed printed course material allowed, personal notes allowed, laptop allowed, access to Internet allowed, pocket calculator allowed. Any form of oral or electronic communication with other students or persons from outside is forbidden. Furthermore, Videos and Screencasts are not allowed.
Module coordinator:	Lorenzo Tanadini (BFH)

Name of the student: *Clara Iglhaut*

School: *ZHAW*

Venue of exam: *online examination upload/download on Moodle*

Declaration of Independent Work

By taking part to this exam, I hereby affirm that the examination is my own work and that I have not used any sources other than those explicitly allowed for the exam. Furthermore, I have not assisted any other students with their online examination.

Exam Briefing

- Write your name and affiliation on the first page
- Next to each problem, the number of points is indicated in parentheses, e.g. (3). Partial credit can be accredited for partially correct answers.
- The level of significance is 5%. Give numeric results (such as p values) to at least three digits.
- Always include a short reasoning (e.g. I applied a marginal F-test and obtained a p value of … , and therefore I conclude ….“)
- **Report all your answers on this document. Convert it as a PDF file before submission.**

Best of luck!

Problem 1

You are involved in a project with the goal to reduce food waste due to improper storing conditions. To avoid food spoilage, your team investigated the results of a controlled experiment in which four storing settings were applied to similar and independent batches of food.

The data set **output_data** contains the data collected in this experiment:

- **output** is the quantity of food that is still edible after storage
- **refrigeration** specifies whether the food was stored in a refrigerated chamber
- **vacuum** indicates whether the food was stored in a vacuum chamber

Set your working directory appropriately and import the data set using:

```
mydata1 <- readRDS("output_data.rds")
```

1. The design of this experiment has a specific name, which one? Is the design balanced? (1)

   ```
   mydata1 = readRDS("output_data.rds")
   str(mydata1)

   with(mydata1, tapply(vacuum, refrigeration, length))
   with(mydata1, table(vacuum, refrigeration))
   ```

   The design is balanced since we have each of the four treatment combinations seven times. We call the design a completely randomized design with factorial treatment structure.

2. Give the **R** code to produce suitable graphical representations of the data set. What do you observe? (2)

   ```
   library(ggplot2)
   ggplot(mydata1, aes(vacuum, output, col=refrigeration)) + geom_point()
   ```

We can see that the food which is stored in the fridge systematically produces higher outputs. We can also see that the food stored in a vacuum seems to produce higher outputs.

3. Give the **R** code to fit a suitable parametric model to this data set. Perform an overall F-test, report its p-value and state your conclusion. (2)

   We fit a factor effect model to the data:

   ```
   mydata1.lm = lm(output~vacuum*refrigeration, data=mydata1)
   summary(mydata1.lm)

   library(car)
   Anova(mydata1.lm, type=2)
   ```

   We get a significant effect for vacuum (p-value = 0.0008651) and for refrigeration (p-value = 2.942e-12), the interaction effect is not significant (p-value = 0.6598518).

4. Assess the model assumptions for your final model: explain what you assess, with which method, give your **R** code, discuss the results and state your conclusions. (3)

   ```
   library(car)
   qqPlot(resid(mydata1.lm))

   plot(mydata1.lm, which=1)
   ```

   I asses the normality assumption visually and with the shapiro there are no visible problems with normality also the shapiro test does not reject the null hypothesis (p-value = 0.6576). Since we do not have that many data points we should not trust the results completely.
   We only asses the equal variance assumption visually, there seem to be no problems with the assumption.

5. Can you simplify the model from 3.? Give your **R** code and state your conclusions (1)

```
mydata1.lm = lm(output~vacuum+refrigeration, data=mydata1)

library(car)
Anova(mydata1.lm, type=2)
```

The overall F-test gave a non significant interaction effect between the two factors, we can fit the model without the interaction effect and only with the two fixed effects for the factors vacuum and refrigeration.

END of problem 1