Exam module 2

It is highly recommended that you answer the exam using Rmarkdown (you can simply use the exam Rmarkdown file as a starting point).

Remember to load the `mosaic` package first:

```
library(mosaic)
```

Part I: Directed reading activities

An educator conducted an experiment to test whether new directed reading activities in the classroom will help elementary school pupils improve some aspects of their reading ability.

She arranged for a third grade class of 21 students to follow these activities for an 8-week period. A control classroom of 23 third graders followed the same curriculum without the activities. At the end of the 8 weeks, all students took a Degree of Reading Power (DRP) test, which measures the aspects of reading ability that the treatment is designed to improve.

Read in the data:

```
reading <- read.delim("https://asta.math.aau.dk/datasets?file=reading.txt")
head(reading)
```

```
##     Group Score
## 1 Treated    24
## 2 Treated    43
## 3 Treated    58
## 4 Treated    71
## 5 Treated    43
## 6 Treated    49
```

Use a boxplot to compare the of measurements of `Score` (the student's DRP score) for `Treated`(direct reading activities) and `Control` visually.

```
## Delete this line and write a command using gf_boxplot(...)
```

Use `favstats` to make a numerical summary of the measurements for `Treated` and `Control`.

```
## Delete this line and write a command using favstats(...)
```

- Write down a point estimate of the mean of the DRP score for students following the new *directed reading activities* and explain how this is calculated.
- Write down a point estimate of the standard deviation of the DRP score for this group and explain how this is calculated.
- Write down a 95% confidence interval for the mean of the DRP score for this group and explain how this is calculated.

Use the command `t.test` to compare the mean of the DRP score of the two groups.

```
## Delete this line and write a command using t.test(...)
```

Go through the details of the output from `t.test`. Your analysis must include an account of

- What the relevant null hypothesis and the corresponding alternative hypothesis is.
- Choice and calculation of test statistic.
- Calculation of p-value and its interpretation in connection to a conclusion of the analysis.
- Calculation and interpretation of a relevant confidence interval.

Part II: Determining sample size

In this part there is no dataset to load into R and analyze. You should just use R as a calculator when you apply the relevant formulas (which are towards the end of the lecture notes for module 1).

To estimate the proportion of Danish companies with less than 10 employees determine the necessary sample size for the estimate to be accurate to within 0.06 with probability 0.90. Based on results from a previous study in 2013, we expect the proportion to be about 0.70.