Experimental Design
Instructions for writing the Report
Due Monday June 26, 2023

- Note that discussion with fellow students during the preparation of the assignment does stimulate statistical and design development type of thinking as well as its implementation. However, the very completion of the report is strictly personal and any evidence of plagiarism will have to be taken seriously.
- Include ONLY key R-output, such as tables, figures and analysis, necessary to answer the question. Any output that is included need to be accompanied by a written interpretation.
- It is not allowed for more than 2 students use the same data set. Might this happen, than this have to be penalized in the marking. To prevent this to happen use the Discussion Forum facility of Nestor to drop a message called "Create Thread" to announce that you are going to use a certain data set.
- Deliver the pdf of your report and a txt file containing R-code via the Report button of the Content page of the Nestor course.

The final assignment consists of performing a project, in which you demonstrate that you are able to apply the most important course content with respect to developing designs and applying statistical analysis of outcomes from real-world experiments. Detailed instructions are provided below.

Write a report consisting of three parts, each focusing on a specific stage of investigation:

1. Part 1: Exploratory or Screening Stage. Find example data for a Taguchi, Plackett-Burman, or Box-Behnken design. Analyse the design as well as the data and draw rational conclusions as far as possible. Make properties of the design explicit. Identify the factors which are most influential on an outcome and try to draw conclusions as far as possible. Discuss pros and cons of your analysis.

2. Part 2: Intermediate Stage. It may very well be better to find another data set for this part. Analyse the type of effects of the identified factors. Distinguish descriptive approaches from formal statistical testing. Use Expert Design software, e.g. R functions from the libraries `FrF2`, `rsm`, or (`bbd`) to design an experiment and to analyse its properties. Give e.g. a thorough 2- or 3-way analysis of variance including interpretation of effects and plotting to illustrate your findings.

3. Part 3: Final Stage. You need to find another data set for this part! Focus on product optimization by a response surface analysis combining two or more outcome variables. Analyse the type of stationary point. Give recommendations with respect to a conceivable industrial productions line. How much certainty is obtained with respect to the uniqueness of the optimum (small prediction error, steep slopes near the optimum). Does robustness of the design play a role? Can you compare with a suboptimal design? What is the prediction interval for the optimal point?

Take the following into consideration:

- For Part 2 and 3 it is useful to add a simulation study on the statistical power or to use software to compute it more directly.
- For each stage make the design properties explicit, such as aliases, resolution, (minimum) abberation, as well as statistical properties such as significance and relevance of effects.
- For each Stage illustrate your Design type of reasoning and statistical analysis with a data example from the literature and/or add constructed data as an illustration.
  - Various references to data are given in the below, but you are free to find interesting data yourself! Think about NIST or data from a scientific article (open science).
  - Indicate how you constructed the data in case if a simulation study, make sure that it reasonably contains some true effects.
    Remember Feynman's saying: " If I cannot construct, I cannot understand."
- Data sets can be found via CRAN its Task View, e.g.
  - libraries with a collection of data-sets: `BHH2`, `EngrExpt`, `agridat`, `daewr`, `dae`
  - libraries with Design software and example data sets: `BsMD`, `DoE.base`, `Vdgraph`, `rsm`, `FrF2`, `OPDOE`, `AlgDesign`, others at `nlme`, `lme`, `lmer`. Yet, others can be found from Task view "Experimental Design" at CRAN.
  - Some example data sets are e.g. `PB12Des`, `PB12Des`, `Reactor.data`, Book exercises such as Excel Problem 12.11
    Note: These are just examples, we expect you to find, describe, and analyse three such data sets, corresponding to Part 1, 2, 3. Find and analyse design type of data from the mechanical engineering field that have your personal interests.
- Your are free in choosing three different data sets even from different contexts, however, do provide sufficient information to follow your writing from an engineering point of view.
- Indicate for each stage for any (statistical) analysis its type: Data descriptive, testing with aliases (confounding), or pure unbiased statistical testing. Always make uncertainties explicit to avoid overconfidence. That is, discuss model fit, model diagnostics, significance and/or statistical power.
- When you construct or use a given design, indicate its formal properties such as resolution, aliases, word length, (minimum) abberation, e.g. in what sense a constructed design is optimal.
- Throughout use Tables and Figures to illustrate your findings and make these independently readable by a fully explanatory caption text. Always provide an interpretation of your results.
- Discuss for each stage resulting recommendations from your design and the findings.

Marking of the report on basis of the following considerations:

1. Part 1: Exploratory Stage:
   Explaining the engineering context of the problem [5]
   Aliasing/confounding [5]
   Statistics/data analysis [10]
   Pros/cons approach [5]
   Conclusions and recommendations [5]
   Hint 1: Add a designs with data analysis example to illustrate the approach.

2. Part 2: Intermediate Stage:
   Engineering context of the problem [5]
   Analyse the type of effects, distinguishing descriptive from formal testing. [10]
   Using Expert Design software to analyse design properties. [10]
   Conclusions and recommendations [5]
   Hint 1: Spent attention to aliases.

3. Part 3: Final Stage:
   Engineering context of the problem [5]
   Analyse the type of effects, distinguishing descriptive from formal testing together with product optimization with correct interpretation of stationary point. [15]
   Paying attention to robustness of factors and (sub)optimality of design. [5]
   Recommendations on conceivable industrial productions line. [5]
   Hint 1: Use a response surface approach.
   Hint 2: A multiple response is preferable.
   Hint 3: Adding a simulation study illustrating properties is beneficial.

Point for free: [10]
Further recommendations:

- It seems very handy to use in each part the sections Introduction, Method, Results, Conclusion & Discussion to order everything.

    1. Introduction: Describe the engineering research problem, it potential implications, and eventually what is currently known.
    2. Method: Give the design of the study and its properties. Do not give any results.
    3. Results: Give descriptive statistics for the most important measurements, the analysis according to the design, the statistical interpretation on testing.
    4. Conclusion & Discussion: Answer the engineering research problem and discuss implications of the findings. Also discuss advantages and disadvantages of the design.

- The idea is indeed that every student analyses and reports on three separate data sets corresponding to the stages mentioned in the above.

- Be creative in finding data and the corresponding engineering context.
- Discuss other possible approaches; give pros and cons.