Evaluation and grading of Forest Biometry and Modelling

Evaluation of Forest Biometry and Modelling does not include a theoretical oral exam. Instead of it, **you will make written assignment** focused on practical statistical skills and its biometrical interpretation. Each of you has an individual data set. Data sets are named "z1" - "z20". You will find the number of your individual dataset in the file "Numbers of assignments". This file itself along with datasets z1-z20 and description of assignments is saved on web page http://user.mendelu.cz/drapela/Forest Biometry/Assignments/.

Grading of your submissions will depend on quality of used R code, correctness and completeness of your results and above all on quality of interpretation of results. Further important evaluated feature of your submission is a **formal and graphical quality of your text** (e.g. reasonable rounding of numbers in results, plots with complete descriptions including descriptions of axes, legends, formatting of text, etc.).

Submission of your works will be provided through "coursework submissions" included in University information system (UIS). Place intended for coursework submissions you can find in UIS on the webpage of Forest Biometry and Modelling ("List of topics and coursework submissions") and it is accessible from any computer connected to Internet. I have already opened coursework submission "room" and you can insert there a file with max. size of to 10 MB. **Please, do not send me your works directly by email.**

Deadline of your electronic submission is January 19, 2025. "Coursework submissions" module automatically saves day and hour of your submission.

Common notes for assignment reports

For all questions submit your answer as well as the R code you used to produce your answer. Integrate output, graphs, and text in a single document - **join both assignments to one document (preferably in Word or PDF)** and **place the R code you used in a separate section at the end of your document**. For R output pasted into your document use a non-proportional (typewriter) font such as `Courier` or `Courier New`. Graphics can be copied and pasted directly from the R studio graphics window, or saved first and then imported into your document.

Place the R code you used in a separate section at the end of your document. Don't include spurious or superfluous code. Ideally I should be able to paste the code you send me directly into my own R console window and get all of the results you report without error messages.

You are allowed to do in groups and help each other, but **everyone has to report your conclusions independently. This means that the included R-codes and graphs may be similar in two reports, but the explanation should be completely your own text. If two students have very similar parts in their reports, then both of these will be rejected.**

Assignment 1 - Basic biometrical analysis

1. Estimate basic statistical characteristics of your data file (for both variables - d (DBH) and h (Height) - mean, standard error of mean, median, mode, standard deviation, variance, coefficient of variance, range, skewness and kurtosis.
2. Estimate confidence interval of mean and standard deviation.
3. Use suitable methods of exploratory data analysis. Use histogram of both variables, add experimental density and density function of normal distribution. Use QQ plot and interpret it. Use box plot and identify outliers (if any). Test normality of both variables.
4. If your sample does not come from normal distribution, use suitable estimate of parameter of location and variability. Evaluate whether your data is suitable for the transformation (is this technique suitable for your data)? If so, calculate transformed mean and its confidence interval.
5. Give short but apt description of both variables (e.g. - Which variable has bigger variability and why? Are there some outliers? If so, what can you do with them and why? What is the interpretation of skewness and kurtosis? Were your samples drawn from a population with a normal distribution? Why we compute confidence intervals and describe their relationship with sample statistics and population ... etc., shortly state all facts you consider to be important for the interpretation of your data)

Assignment 2 - Choice of suitable biometrical model

Analyze your data (X variable - d - DBH, Y variable (response) - h - Height.

A) Fit the **linear regression model** ($y = a + bx$).

- Write your specific model (including calculated parameters) and decide whether this model is statistically significant (and what it means from practical point of view) and why.
- Write all measures of correlation (coef. of determination and coef. of correlation) and describe their meaning and importance in analysis.
- Assess significance of individual parameters and, if some parameters are insignificant, re-write model into a suitable formula.
- What values can regression parameters in population reach?
- Use basic methods of regression diagnostics (analysis of residuals, identification of outliers, if any, heteroskedasticity, etc.) and evaluate quality of model and data.

B) Compute **suitable nonlinear model** for the same data and index of correlation, its significance, confidence interval of parameters and basic regression diagnostics.

- C) Compare both models and decide which one is better from a statistical and biometrical point of view.