**Slide 18**: First, solve the problem. Then, compare your solution to mine. Take really as much time as you need, it is important to be solved correctly. Problem: I have data from a normal distribution with a mean of 28 cm and a standard deviation of 3 cm. I need to find the boundaries for 30% of the data closest to the mean. Is my solution correct? qnorm(0.3, 28, 3). Give me 1) found values in csv, 2) your solution as R code, 3) plot for this normal distribution here in chat, 4) print this plot to pdf

**Slide 21**: Write for me the R code for this situation: I am working with the dataset [insert table name]. **Create boxplots using the boxplot()** function for [name of the column of the quantitative variable] divided by [name of the column of the qualitative variable], which has a **number of levels** [insert the number of levels of the qualitative variable]. Individual levels are **labeled** "X1" [insert level name according to dataset] for "Y1" [insert level name reflected in the graph], "X2" for "Y2" [continue according to the number of levels]. Make the boxplot so that it is/is not [choose one option] **horizontal**, with/without [choose one option] **notches**, with the **color** "Z1" [choose a color] for "Y1" [insert level name reflected in the graph], with the color "Z2" for "Y1" [continue according to the number of levels]. The **description of the x-axis** [y for horizontal graph] will be "" [insert your own name for the qualitative variable], the **description of the y-axis** [x for horizontal graph] will be ""[insert your own name for the quantitative variable], **the labels on the y-axis will be horizontal/vertical** [choose one option]. The naming of the levels of the qualitative variable will be "Y1" for "X1", "Y2" for "X2" [replace according to the key as above].

**Slide 22**: Step-by-step leveraging external tools, analyze in R and fill the answers in this template: **After conducting Welch's t-test, we can say that there exists/is no significant difference in** [insert reasonable name of dependent variable] **between** [insert the name of the first level of explanatory variable] **and** [insert the name of the second level of explanatory variable] (df = [degrees of freedom], t = [t-value rounded to two decimals], P = [p-value rounded to three decimals]).

**Slide 26**: <https://chat.openai.com/g/>[g-ChD77CFZ8-tutor](https://chat.openai.com/g/g-ChD77CFZ8-tutor)

**Slide 29**: **Using my notes in Czech, use the creative writing technique to write a data analysis methodology in English in the style of a scientific paper methodology. Notes**: Statistická analýza v R 4.2.2, nejdříve "vegan", abychom spočetli tax diversity mechových společenstev. Konkrétně nejdříve taxa2dist (napiš co to je ne funkci) a pak taxondive (napiš co to je, ne funkci), pak to bylo použito spolu s richness a coverage mechů v GLMMs (dej plný název). Vysvětlující proměnné: typ skály podle těžkosti lezení (Control, High difficulty, low dif), a vertikální gradient (od 0 do 12 m), zvážili jsme i jejich interakce a taky alternativní verze modelu s polynomy 2. s. Final model vybrán AIC. Random term byla specifická skála, kde se ten vertikální gradient měřil. Konkrétně pro coverage jsme použili glmmTMB balík, model s beta distribucí. Pro richness "lme4" s neg bin rozdělením z MASS. Taxonomic diversity → glmm s gama dis. Pak anova z car balíku, poté pro zjištění pair-wise rozdílů mezi typy skal emmeans balík, napiš jaká se tam používá defaultně adjustace p-value. Grafy pomocí interactions a ggplot2.

**Slide 36**: Take a deep breath and work on solving the problem step by step. I need to build an LMM model. What the R code would look like for the model I need. Here is the research design: Individual volunteers "Volunteer\_ID", belonging to three age categories "Volunteer\_agecat" underwent three types of tests "Test\_type", 5 times each type of test, i.e. 15 attempts in total, and it was observed how long it took them to complete the test "Complet\_time". Because of the possibility that volunteers may improve over time, tests are presented to volunteers in a scheduled order "Test\_order". Rules for building the model: A specific "Volunteer\_ID" belongs to only one "Volunteer\_agecat", we are not interested in the influence of either one. All "Test\_order" and "Test\_type" levels apply to each "Volunteer\_ID". The only thing we are interested in is the influence of "Test\_type" on "Complet\_time", we want to suppress everything else, but we have to take into account that the factors are crossed and nested differently, see the description above.

**Slide 47**: Read the following research and build an LMM model in R accordingly. Here is the research design: Individual volunteers "Volunteer\_ID", belonging to three age categories "Volunteer\_agecat" underwent three types of tests "Test\_type", 5 times each type of test, i.e. 15 attempts in total, and it was observed how long it took them to complete the test "Complet\_time". Because of the possibility that volunteers may improve over time, tests are presented to volunteers in a scheduled order "Test\_order". Rules for building the model: A specific "Volunteer\_ID" belongs to only one "Volunteer\_agecat", we are not interested in the influence of either one. All "Test\_order" and "Test\_type" levels apply to each "Volunteer\_ID". The only thing we are interested in is the influence of "Test\_type" on "Complet\_time", we want to suppress everything else, but we have to take into account that the factors are crossed and nested differently, see the description above. Show me the code for LMM with special emphasis on …

**Slide 49**: Imagine three different experts are answering this question, all experts will write down 1 step of their thinking, then share it with the group. Then all experts will go on to the next step, etc. If any expert realises they're wrong at any point, then they leave.   
The question is...