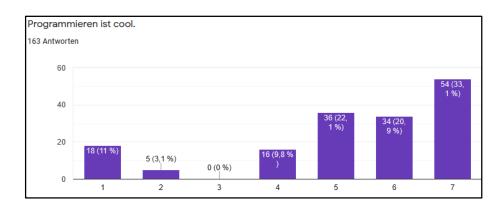
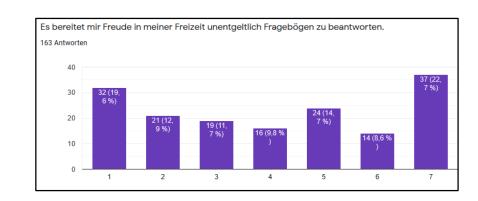


Everything combined

FINAL ASSIGNMENT







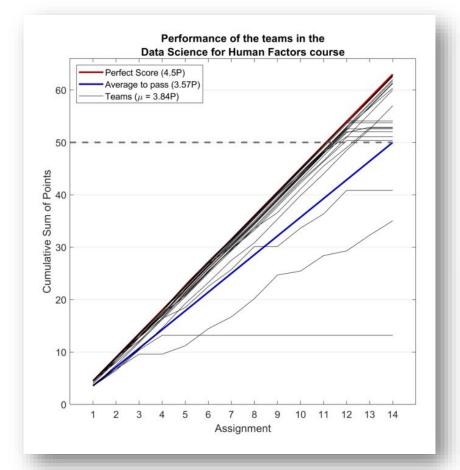


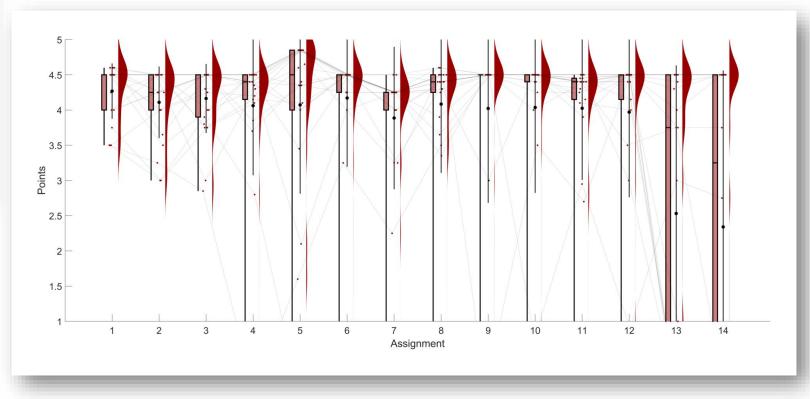
Topics

- Introduction: GUI and basic calculations
- Coding 1: Scripts, style, and variable classes
- Coding 2: Control statements and loops
- Visualization 1: Basics, subplots, get and set
- Coding 3: Functions
- Visualization 2: Descriptive plots
- Coding 4: Basic input and output
- Visualization 3: Distribution and 3D plots
- Coding 5: Input and output specials last lecture before holidays
- Machine Learning 1: Introduction and dimension reduction
- Machine Learning 2: Clustering
- Machine Learning 3: Classification
- Coding 6: Efficiency and debugging basics
- Coding 7: Advanced functions and debugging













The final assignment

- NOT in teams any more! You are on your own!
- Deadline: 1 month from today! Enter grades in QUISPOS must be end of this semester... it might be possible to enter grades later but that must be done manually via the Prüfungsamt. If you need a deadline extension, let me know in advance!
- The idea of this task is to combine all previous tasks
- You know the "Kleiner unspezifischer Fragebogen" which you hopefully filled in (and sent around).
- An EXCEL sheet with the answers will be available for download on ISIS
- This questionnaire should now be analysed completely
- You are quite free in how exactly you want to do this, just science the data!
 - Create a script that does the complete analysis (including running all subfunctions)
 - Comment which MATLAB version was used
- No specific naming convention necessary
- A few things are required, though, to show you mastered the course...





What I expect in the script(s)

- Load the EXCEL sheet into MATLAB
- Clean the data (remove all data of 2019)
- Bring it in a usable form (some answers are numbers, others are text)
- Use at least one for or while loop
- Write a function and use it at least twice, choose sensibly why this should be done in a function.
 - Remember the help and input checks
- Visualize the data as you see fit
 - All final plots have to have appropriate size, font size, axes, legends where due etc. In short, they should be real figures which are usable and you and I can show around.
 - Use subplots at least once. Other fancy plots are of course possible as well, but the main point is that the plots should fit the data.
- Statistics are not a necessity, but correlations and t-tests could be a nice addition if you want to make a point (the functions are: corr, ttest)
 - This is not part of the course and thus won't be graded but I will see and like it.
- Concerning the ML part:
 - Use tsne at least once
 - Use clustering at least once
 - Explain whether or not the clustering makes sense in the case of this data and why you think that
 - Use classification at least once
 - There are many two-classes questions which might be predicted by the numerical questions
 - Examine whether your classification exceeds statistical chance at the alpha = 0.05 level
 - Make sense of your results, plot them in a meaningful way
- Export your final figures, they should be self-explanatory
 - Movies/gifs don't have to be used. If you find a nice way to do it, it's great though, of course!
- Create a zip file with your entire script and all nice exported plots. If it exceeds the ISIS data limit export the plots only as tif or jpeg, or hand in a link to another file hoster, or stop by my desk and hand it in manually.



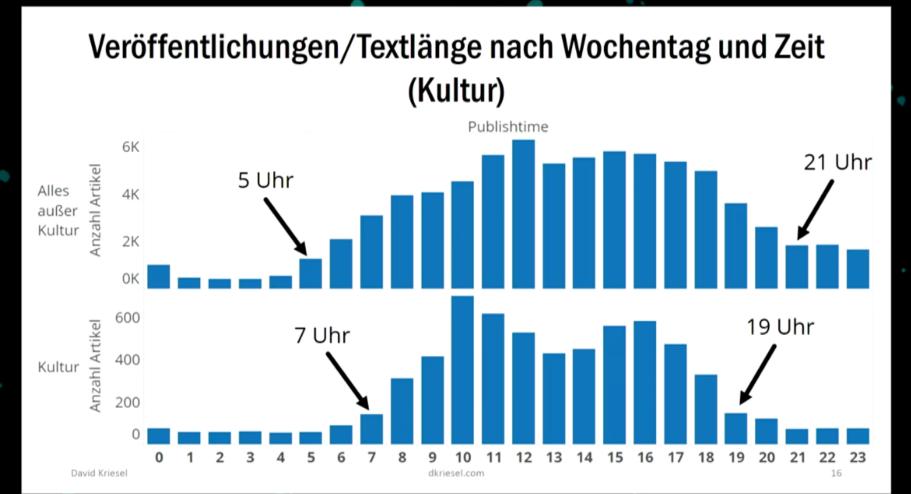


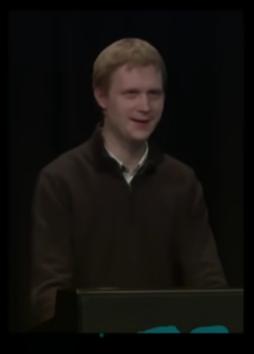
Last but not least

- Comment every step that is non-trivial
 - Explain your reasoning
- Use clean code, sensible variable names, sensible sections
 - Make sure I, and any other person reading the code, can understand easily what is going on.
- In short: Code well, not fast.
- Have fun © I think the questionnaire is quite interesting and I'm sure you will find things you are curious about in the data!











4 / 58:45





SpiegelMining -