Mandatory assignment 2

Please submit:

1. A report in PDF

2. In a zip file: the code in .R, .rmd, or .qmd and the data used (you can use the data you used in mandatory exercise 1.)

Please write comments in the code and reference the lecture where you found it. If you used a language model, please include your prompt in the appendix.

Try to work together and learn from each other. We will give feedback for the full report, and that is more useful, if you did try all tasks. **Only use ChatGPT/Copilot for assignment to give feedback on your solution.** If you don't know the solution – write instead, that you don't know how. Then I will know where I have "failed" in my teaching. Don't use it to write your code, as you will learn more from trying to find the answer yourself. If you use it, remember to write what you used it for.

You should work with the data from last lecture – or you can find some new if it is not appropriate. You can ask.

Tasks:

- Load the dataset into R

- Create a point plot or jitter plot.
- Create a boxplot or a bar plot.
- Generate a histogram to visualize the distribution of one of the numerical variables.
- Create a visualization of own choice that you have "polished", and write what the purpose is with your visualization. (Is it for a managerial decision? A publication? How?)

- Create a linear model in which you model using at least tree input/independant variables.
- Visualize the model output using dotwhiskers and stargazer
- Reflect on the model – how could it be improved?
- Reflect on how you have used Generative AI (if you have).

If you are uncertain about what could be input/independant variables and response variables, feel free to reach out.