

YOUR TAKE HOME EXERCISE: BASIC DATA HANDLING WITH LLMS

- Use an own dataset (no confidential data!) or a publicly available data set, like this from Kaggle:
<https://www.kaggle.com/datasets/parve05/customer-review-dataset>
- With easy prompting, try to achieve the following:
 - Define a plan for analysis
 - Clean data
 - Identify and visualize statistics, such as means, standard deviations, group comparisons...
 - More advanced: Analysis of longitudinal data, such as the correlation between review length and helpfulness
- Document your prompts and results in a one-pager / slide
- Send the results via mail with a short note if I may use them as **examples** in the slides
- For yourself: What do you think are the problems with this approach? Which hurdles did you run into?