HW2

Alex Pasternakova

9.9. 2021

1 Q1: Does OpenRefine alter the raw data during sorting and filtering?

No, it stores raw data as well as steps in code of how the copy was altered/cleaned up for research purposes. The raw data is always stored.

2 Q2: Fix the interviews dataset in OpenRefine enough to answer this question: "Which two months are reported as the most water-deprived/driest by the interviewed farmer households?"

October and September were two months mentioned the most by farmers.

3 Q3: What were the 10 most frequent occupations (erhverv) among unmarried men and women in 1801 Aarhus?

Using OpenRefine I tried to clean up the data as much as possible, however, I don't speak Danish and turns out google doesn't either... So here is my guess:

1. soldat (includes landsolat, national soldat etc) 2. tienestepige/tienestekarl 3. væver 4. vanfør 5. nyder almisse af sognet 6. inderste (this keeps being mentioned but translation only says "inner", I ponder the idea of it not being a real occupation) 7. læredreng 8. indsidder 9. skræder 10. hospitalslem

OpenRefine script LINK