

W45: Open Refine

1. Create a **tidy** spreadsheet/table listing the names of Danish monarchs with their birth- and death-date and duration of reign. They should be sortable by year of birth. Suitable source websites are [here](#) and [here](#), but you can also use another source, provided you reference it. (Collaboration is welcome. Remember to attach this spreadsheet to Brightspace submission)

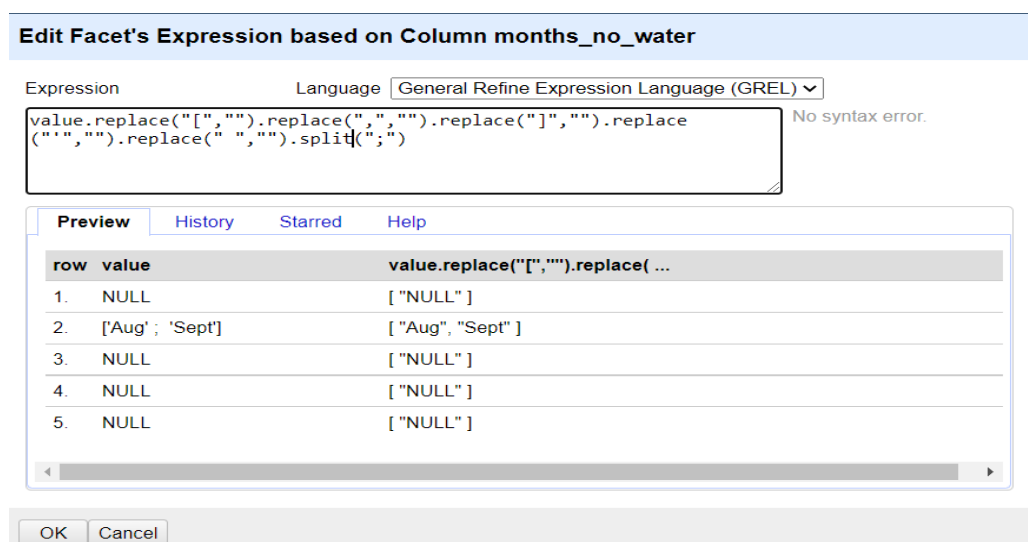
The spreadsheet is attached

2. Does OpenRefine alter the raw data during sorting and filtering?

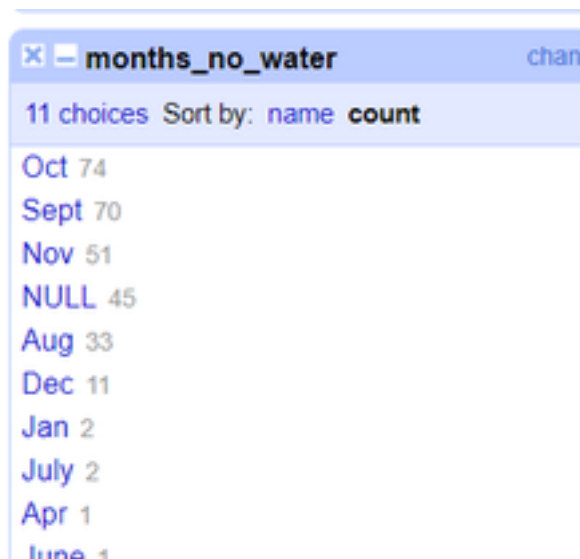
No. The reason we use OpenRefine is to process data, not create it. The benefits of OpenRefine is the fact that we can look at the entire editing history, and in that way, ones conclusion can be recreated by others.

3. 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?"

In order to figure this out I wrote the following code in OpenRefine in order to remove excess characters, and split the months up separately in the column "months_no_water".



After doing this we can see that the driest months are October, which is mentioned 74 times, and September, which is mentioned 70 times. This is the complete result



months_no_water	
name	count
Oct	74
Sept	70
Nov	51
NULL	45
Aug	33
Dec	11
Jan	2
July	2
Apr	1
June	1