**Home Work by Christian Kjeldsen**

**1**

Part one sort and file the danish monarchs by Regin and death, and date of birth in a excel spreadsheet.

In this exercise I created one column for each reign king or period with one or more kings and the ones without a king.

I then created a column for the year of the start of the reign, and one for the death of the king or end of reign.

I then noted the year of birth for each year and for those kings where this data was absent, I noted N/A for no data available this I did in a separate column. To see this data place open the excel file “Kongerne V5” I used the data both from the royal family themselves and the university’s list

**2 try to sort the data in open refine**

I now took the data I plotted in excel and tried to restore it in Open Refine, with no altering of the data when the change of format occurring but you need to mark that you want to see the entire data set and not just 25

To see the data format from the Open Refine <http://127.0.0.1:3333/project?project=2251572188600>

I now tried to sort the column marked reign “ended in years” so that it now started with the last year and not the first as I had done in excel and in the same excel document, I had made the 4. Columns depend on the 2. And 3. Column but to my surprise the 2. And 4. Colum also started with the latest so there was no change in the calculated period of reign

So it semes that the program were able to see that the horizontal data where connected it did not alter the 4. Column to my surprise.

**3**

**I needed to find the two moths who had been the driest in the report from the farmers**

Here we were given a set of data In a excel sheet where I first needed to identify the right column for the declaration off no water in a specific month reported by a farmer.

I found the column AG in the excel sheet then I copied that column to Regex.

Here I inserted the data from column AG in the Test String.

Then I used the regular expression [A] to mark every Aug for August. declaring the farmers who had marked that month as a month with no water.

The total amount of farmers who reported August as as a month with no water was 33.

(Be aware that April is also marked by this but it is only present In one cell)

I then used the same approach for the other months with the First letter for every month inserted in the regular expression

September [S]= 70

October [O]=74

November [ov] this because using N would have counted the Null values as well and November was marked by Nov so [ov] caught that month. Resulting in a total of 102 hits

May [M]=1

April =1

Marking the letter [J] results in 5 hits where 2 is January and 2 is July and 1 hit for June

The last month listed is December where I have used the description [D] resulting In 11 hits

Now summing these results we can conclude that the month with the least water according to the farmers was November with a hit rate of 102 reporting of “No Water”

And the 2. Most reported month that had “No Water According to the farmers was October with a total of 74 famers reporting of “ No Water

Here is the regex for every approach

December: <https://regex101.com/r/vGJGQL/1>

August: (counting April as one result as well) <https://regex101.com/r/6OCDHZ/1>

October: <https://regex101.com/r/HlCy85/1>

September: <https://regex101.com/r/Ein6d4/1>

November: <https://regex101.com/r/deYfaP/1>

January, July, June: <https://regex101.com/r/rttQEE/1>

May <https://regex101.com/r/xfQvZZ/1>