1:W35: Regular expressions and spreadsheets

DESCRIPTION
Upload your answers/solutions to the problems below. Beware of making the submission legible and understandable to another reader:

1. What regular expressions do you use to extract all the dates in this blurb: http://bit.ly/regexexercise2 and to put them into the following format YYYY-MM-DD ?

    (\d+).(\d+).\s?(\d+)   ← to make each part of the date it's own variable

    $3 - $1 - $2   ← to put the variables in the desired order

2. Write a regular expression to convert the stopwordlist (list of most frequent Danish words) from Voyant in http://bit.ly/regexexercise3 into a neat stopword list for R (which comprises "words" separated by commas, such as http://bit.ly/regexexercise4). Then take the stopwordlist from R http://bit.ly/regexexercise4 and convert it into a Voyant list (words on separate line without interpunction)

    Converting stopwordlist to a neat stopword list for R

    Regular expression: (\S+)\n

    Substitution: "$1",

    Converting a stopword list for R into a Voyant list

    Regular expression: (", ")

    Substitution: \n

3. In 250 words, answer the following question: "What are the basic principles for using spreadsheets for good data organization?"

    Data validation, in order to control how and what data is entered. This includes drop down menus or only allowing certain numbers/range of numbers to be entered to minimize erroneous entries. Only having one value in each box, one entry per row and one variable per column. Remember that changes made cannot be tracked or undone easily so leave the raw data alone and work off another version of the data. Colors won't be transferred when exporting elsewhere so you can use them but without a legend or explanation it can be confusing later.

4. Challenge (OPTIONAL)!Can you find all the instances of 'Dis Manibus' invocation in the EDH inscriptions in https://bit.ly/regexexercise5? Beware of the six possible canonical versions of the Dis Manibus formula (see day 1 slides)!

Work in progress…

\b[Dd]i. [Mm]anibus

\b(di*s? manibus) (sacrum)?|dms?

(\bd m s?)|(\bdi*s? manibus)|(\bdi*s? manibus sacrum)
\bdi*s? ?m(anibus)? ?s?(acrum)?