**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 ?

To find all dates: (\d+).(\d+)..(\d+)

To substitute: $3 $1 $2

1. 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)

First (")|(,) replace all “ and , with nothing.

Take the text and find all white spaces \s and replace with \n

To get back.

\W replace “, “

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

Have a data documentation sheet that describes each variable, its variable type, what does the variable represent i.e., what does it measure. Be concise and have a standardized way to do your data documentation so it is easily readable. Tidy format i.e., each row represents an observation, and each column corresponds to a distinct variable. Encode your variables correctly, binary, and categorical variables as factors, integers as discrete and floats as continuous etc. Remember to document whether your data is in a wide or long format. You can change these formats by functions such as pivot\_longer or pivot\_wider. Code your missing values in the same manner i.e., don’t change style. Check if dates are coded uniformly or whether they are in different formats.

1. 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)!