1:W35: Regular expressions and spreadsheets

1. Regular expression: (\d+).(\d+).\s?(\d+)

Substitution: $3-$1-$2

1. Stopwordlist
   1. From Voyant to R:
      1. Regular expression: (\S+)\n?
      2. Substitution: "$1",
   2. From R to Voyant:
      1. Regular expression: "(\S+)",?
      2. Substitution: $1\n
   3. It is better to make a lot of columns for the different variables instead of making three different tables for different aspects of the data (e.g. Plot, Livestock etc. – don’t put them into separate tables, but put them all into one, if possible.
   4. Be consistent with variable names, i.e. don’t make typos, so program thinks that Raucu and Rucu are two different things, and not just an error.
   5. When making names for the columns, be sure to be consistent in the convention used, e.g. when making the variable names, then never use a regular space, but join the words with an underscore (i.e. \_ ). It can also be a good idea to consistently use lowercase letters for all variables.
   6. When exporting data from a spreadsheet, for example from a Microsoft Excel sheet, be aware that it is only the data from the currently selected sheet that is going to be exported.
   7. Be sure to be consistent and logical when choosing names for files that will facilitate the use of regular expressions to call forth the files.
   8. Follow the convention of writing dates in the following format: YYYY-MM-DD
   9. Be sure to only include one value for each data cell (combination of row and column in a data frame) – if there is an instance of several values in a cell/column, simply create a new column to assign the excess value to.
   10. In spreadsheets, do not use color coding to code information into the data, for example to signify the importance of the data. The reason to this, is that the information will get lost when exporting the data to a csv file, since it will be not readable in a meaningful way.
   11. When working with others while collecting data, be sure that you follow the same exact format ­– this can be achieved by employing data validation, which will help avoid errors that can arise when typing in data points.
   12. Always, always make sure to have a backup of your data and preferably to have it stored in more than one place. Even though cloud storage offers accessible and inexpensive storage options, it might be a good idea to have a backup on a local, physical drive, in the case something went wrong.