Homeworks week 36

Regular Expressions

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 ?

This is the link to my response:

<https://regex101.com/r/wtlnsK/1>

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)

Task 2 has been solved in this link:

<https://regex101.com/r/DkIPwx/1>

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

Spreadsheets are often used as a versatile tool for data entry, storage, analysis, and visualization. In this course, we use spreadsheets for data entry and data storage.

Using spreadsheets for effective data organization requires adherence to several fundamental principles. This is because ensuring the accuracy of data user-friendly, thereby reducing potential future errors.

One important thing to remember when working with spreadsheets is entering dates. It’s important to use the global standard: [YYYY-MM-DD]. Writing dates in other formats can create conflicts and may cause issues with the data entered in spreadsheets.

Furthermore, to minimize data-related problems, it’s crucial to maintain consistent formatting for all data entries. This also involves avoiding merged cells, as it can complicate sorting, filtering, and calculations. Each cell should contain a single unit of data.

Another critical aspect of working with spreadsheets is data protection. Keeping track of spreadsheets versions to monitor changes and revering to previous states when necessary is essential. It also involves designing your spreadsheet to accommodate potential growth in data volume, thereby avoiding embedding values whenever possible.

These are just some of the basic principles to remember when working with a well-structured spreadsheet where you can efficiently organize and store data. It also helps reduce the risk of errors and confusion.