**Learning Journal**

**October 30th:**

We started out with exercise 3, doodling around with it on regex101.com. The end goal is to try and make the list in alphabetical order. In order for us to create a “neat” stop word list, we decided to try and remove all of the quotation marks first. That is done with

[“.”]

Then the text that appears in ‘’substitution” is copy-pasted into “text string” - this way all the quotation marks are removed from the original text. Next order of business is to remove the commas, and working around with it has taught us, that

[,]

will suffice, meaning

[“]

would have been sufficient in the first step. A lesson has been learned. Then the text that appears in ‘’substitution” box is copy-pasted into “text string”. Now there are no quotation marks and no commas in the text. Then we started the process of alphabetizing the list, and our idea of it was that we wanted to find all of the words that started with an “a”. We tried that using the following command:

^[Aa]

This, however, did not work as intended, though by mere chance we discovered that using

\s[Aa]

This only works, because the first word in the search list start with “SPACE a”, but for our purposes it works just fine. At this point we had found all the words starting with “a”, but did not know what to do about finding the whole word. We consulted the list of commands on librarycarpentry.org, and because of this we were able to construct the following command:

(\s[A-a]\w\*)

which grabs the entire word starting with an “a” in the stop word list. Using

(\s[A-a]\w)

it only grabs us the first part of the words starting with “a”, therefore the “\*” is essential, because it allows us to mark the entire word. During this exercise it had bothered us, that we did not know how to write several commands in the “regular expression” box, but it turns out that it can be done using this command “|” between commands. With this in “tool” in our toolkit we were able to start grouping the words in the stop word list, by going

(\s[A-a]\w\*)|(\s[B-b]\w\*)|(\s[C-c]\w\*)

and so forth. Doing this we can group all of the words starting with the same letter together. At this point in time we discovered a huge flaw in our way of doing this work: The program cannot read Danish special letters (“æ”, “ø” and “å”), meaning if a Danish special letter appears within a word, the program does not grab the rest of the word. We did not know how to make up for this mistake so we powered on. At this point we had worked on this for an hour and a half, and decided to work on something else.

**November 2-9:**

During this entire period, I have been extremely ill with the flu. This has put me behind in every class, not least in this class. I went to the lecture Monday the 4th, but not the class the next day, as I had a high fever. I have a lot of catching up to do.

**November 12:**

Today in class we installed gitbash, which is a program that allows us to use our shell and run alongside it. I have a Mac PC, so I was already able to open my shell (called terminal), but gitbash adds commands and functions to it. It was a very new experience to use a shell, at least for me - but it turns out that others in class had the same experience. It was therefore regrettable that I did not end up grasping the full potential of the program, and I have a hard time picturing how I will manage to find a use for it in the final exam paper. The class had a high tempo, but I can only blame myself for not figuring out the program’s potential.

**November 12 - later in the day:**

It turned out that I wasn’t the only one who had problems with gitbash, and therefore we were a couple of guys from the class who sat together, and worked on the exercises provided by the course on <https://swcarpentry.github.io/shell-novice/>. It turns out that the work helped me understand a little more of what my terminal can be used for, and I got to be familiar with a couple of commands. It was very useful to continually use

man ls

which brings up the manual. That way a lot of the questions that firstly appears can be answered. I also got somewhat familiar with using

mkdir

which makes directories directly in the terminal. When that happened, it was also a challenge for me to keep track of the directory I was in, but ended up being able to figure out maneuvering back a layer using

..

I must, however, admit that I have a very hard time imagining myself ending up using gitbash on my terminal for my final paper in this class. I would have to become much more familiar with it, and use a much higher amount of data, than I intend, for it to become a useful tool, the way I see it.

**November 18:**

Class of today started out with some clarification on how to find the data for the final paper, and what to do with it. That took away some anxieties and cleared up some of the questions I have. I am still - however - a little unaware of the extend of the paper, which could be because of the class I missed due to illness. I must figure that out at some point, by asking some of the other students in class at some point. Adela also made a comforting point about the fact that even specialists in this field gets frustrated with data management.

**November 19:**

Before class today we had to download R and R-studio which, as I understood it, was to be the culmination of the work over the last few weeks. I mean in the sense that it appeared to me that all the confusing and frustrating information and tools we have been introduced to, was to be the basis of learning how to work with R. Therefore, I went to class today with a mixture of anxiousness and anticipation.

The course started with an introduction to the program, where we were told how to use basic commands. This included being able to create folders and paths between them. Later on me moved on to some more advanced commands - I especially found the vector grouping interesting and difficult, because it seemed to be appropriate for messy or incoherent data sets.

**November 25:**

In the morning of November 25th, I started looking at the exercises regarding R Studio. I installed the tidyverse package using the following command:

install.packages("tidyverse")

That felt like running a “real” command, giving off the impression that I put in a line of code, and stuff started to happen. Afterwards we wanted to start doing the exercises called “Introduction to R”. We realized that the work was done during class November 19th, working around with objects and variables - assigning them values and playing around with that. What we did on the morning of 25th was running through it again a couple of guys from class. We discussed it among ourselves and reviewed the work. I feel close to comfortable with assigning values using the

🡨

command.

**Lecture:**  
Network theory.

Small world theory. What is the distance between two random people? It turn out that to people can find each other with only 6 levels of separations.

Strong / weak ties. Most people get jobs through weak social ties. This might contribute to a more dynamic network

If two people are strongly connected. They properly have the same knowledge and perspective. This might not be ideal in the workspace

In a network the more interconnected entities are the most influential.