

My script to Jupyter Lab - the text will referate to the number of the column in the ipynb file
Jupiter Lab work with coding language Python

The script for the DRB_1367_1370.ipynb file

3. First I install dacy in Jupyter Lab
4. After that I import spacy, which is the english version of dacy, then I import dacy and finally I import pandas as pd, because it is easier to write
5. Then I use the nlp package to load the large model in dacy
5. I use the pwd function to see my workspace
6. In this column I load the dataset into the program using pandas and the read.csv function
7. I tell the program to use the nlp package on the datasets column "Tekst"
8. The program is know searching for persons and the docid in the "Tekst"
9. I run the "per" to view all the names on persons that it found
10. I create an empty dataframe and name it "DRB_1367_1370_name"
11. Then it tell it to make 2 columns in the dataframe called "docid" and "name" for the "docid" and the "per"
12. I run "DRB_1367_1370_name and see the dataframe
13. At last I convert the dataframe to an csv file, so that I can use it R and make a visualization

The script for the DRB_1371_1400.ipynb file

1. First I import spacy, which is the english version of dacy, then I import dacy and finally I import pandas as pd, because it is easier to write
2. Then I use the nlp package to load the large model in dacy
4. In this column I load the dataset into the program using pandas and the read.csv function
6. Here I tell the program that the "Tekst" is a string and not a number as it assume
7. I tell the program to use the nlp package on the datasets column "Tekst"
8. The program is know searching for persons and the docid in the "Tekst"
9. I run the "per" to view all the names on persons that it found
10. I create an empty dataframe and name it "DRB_1371_1400_name"
11. Then it tell it to make 2 columns in the dataframe called "docid" and "name" for the "docid" and the "per"
12. I run "DRB_1371_1400_name and see the dataframe
13. At last I convert the dataframe to an csv file, so that I can use it R and make a visualization

I have used this website:

<https://towardsdatascience.com/structured-natural-language-processing-with-pandas-and-spacy-7089e66d2b10>