

Prosjekt

EIRA

2022-07-04

Packages

Download data - Since the data we need and use for our assignment is provided by NORAD themselves, we used their website to download the dataset. - From the website we chose the variables needed and for the preferred years. We would like to include all variables and from the years 2013 til 2021, since those were the years that NORAD used. - We also made the data into a dataset. - We also made the variable names more tidy, and had a quick look at the dataset.

```
Norad_Norwegian_development_assistance <- read_csv("~/Documents/Statsvitenskap /Political Data Science 1  
Norad_clean <- Norad_Norwegian_development_assistance %>%  
  clean_names()  
  
glimpse(Norad_clean)
```

Creating a sub dataset After having a brief look at the data and carefully choosing the variables needed, we made a smaller dataset for it to be easier to work with.

```
NORAD_subset <- Norad_clean %>%  
  select(agreement_number, description_of_agreement, sdg_focus)
```

TF-IDF - This is where the real work starts.

```
tfidf1 <- NORAD_subset %>%  
  count(description_of_agreement, name = "ord") %>%  
  bind_tf_idf(agreement_number, description_of_agreement, ord)
```