PolNos: Political Nostalgia in Party Manifestos

Version 1.0

Authors: Stefan Müller and Sven-Oliver Proksch

Table of Contents

1 About *PolNos* **1**

2 Manifesto-Level Dataset **2**

3 Sentence-Level Datasets **5**
3.1 *PolNos* Sentence-Level Data 5
3.2 Hand-Coded Manifesto Sentences 5
3.3 Hand-Coded Manifesto Sentences for Comparison of Measures 6
3.4 Full Machine-Translated Corpus 6

4 Nostalgia Dictionary **7**
4.1 PolNos Classifiers 10

5 Load *PolNos* Datasets **10**

6 Merge *PolNos* with *Manifesto Project Dataset* **10**
6.1 Load and Prepare *Manifesto Project Dataset* 11
6.2 Merge Manifesto-level *PolNos* Dataset with *Manifesto Project Dataset* 11
6.3 Merge Sentence-level *PolNos* Dataset with *Manifesto Project Dataset* 12
References 13

1 About *PolNos*

PolNos is a collection of datasets on political nostalgia in party manifestos. The current version of *PolNos* contains 1,648 manifestos from 24 countries, released by 379 parties between 1946 and 1946. The main datasets come in two formats: either on the level of (quasi-)sentences or aggregated to the level of manifestos.

In this document, we describe how to load and use the *PolNos* datasets. We provide an overview of all variables included in five datasets: *PolNos Manifesto-Level Dataset* (Section 2 and Table 1); *PolNos Sentence-Level Data* (Section 3.1 and Table 2); *Hand-Coded Manifesto Sentences for Training Classifiers* (Section 3.2 and Table 3); *Hand-Coded Manifesto Sentences for Comparison of Measures* (Section 3.3 and Table 4); *Full Machine-Translated Corpus* (Section 3.4 and Table 5). We also provide the nostalgia dictionary in various file formats (Section 4).

In Section 5, we show how to merge the PolNos data with datasets and corpora released by the *Manifesto Project* (Volkens et al. 2021).

> **Note**
>
> If you use PolNos data in your own research, please cite the accompanying paper and PolNos data repository:
>
> Müller, Stefan and Sven-Oliver Proksch. 2023. *PolNos: Political Nostalgia in Party Manifestos.* Harvard Dataverse, V1. DOI: 10.7910/DVN/L198GI.
>
> Müller, Stefan and Sven-Oliver Proksch. 2023. Nostalgia in European Party Politics: A Text-Based Measurement Approach. *British Journal of Political Science* online first. DOI: 10.1017/S0007123423000571

2 Manifesto-Level Dataset

We provide a *PolNos* dataset aggregated to the level of manifestos. The file is called `data_polnos_manifestolevel` and is available in `rds` and `csv` file formats. This dataset contains one observation per manifesto and includes the aggregated measures of nostalgic rhetoric for the six measures introduced and validated in the accompanying paper.

If you use variables from the *Manifesto Project*, please cite Volkens et al. (2021). Variables on government compositions come from *ParlGov* (Döring and Manow 2021) and were added manually for missing parties. Economic variables were included in the *Comparative Political Data Set* (Armingeon, Engler, and Leemann 2022). Please cite these datasets in addition to *PolNos* if you use these variables in your analysis. Table 1 describes all variables included in the manifesto-level dataset.

The dataset `data_polnos_manifestolevel` consists of 1,648 observations. Each observation is one party manifesto. Figure 1 summarises the countries and elections included in the dataset. The numbers beside country names show the number of manifestos and elections.

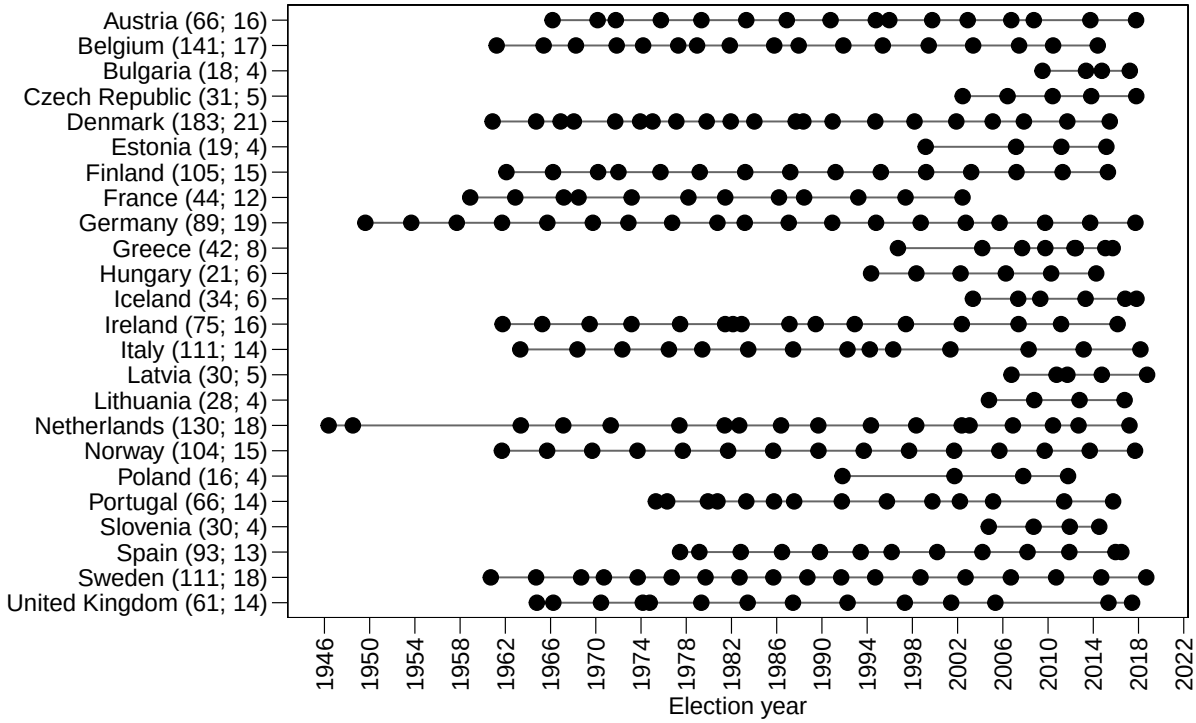

Figure 1: Available party manifestos in PolNos dataset (v 1.0)

Table 1: Variable description: Manifesto-level dataset

Variable	Details
nostalgia_main	Main meausure of nostalgia (DistilBERT), same as nostalgia_sentences_per_1000_bert
election_id	ID for each election
manifesto_id	ID for each manifesto (party ID + election date)
year	Election year
edate	Date of election (format: YYYY-MM-DD)
date	Date identifier (format: YYYYMM)
party	Party identifier (from Manifesto Project)
parfam	Party family coding from Manifesto Project
partyname	English translation of party name
partyabbrev	Party abbreviation
party_family_recoded	Party family (recoded into eight categories)
countryname	Country name
region	Region (Northern, Western, Southern, or Central- and Eastern Europe
unemp	Unemployment (based on CPDS)
unemp_lag1	Unemployment, lagged by one year (based on CPDS)
unemp_lag2	Unemployment, lagged by two years (based on CPDS)
unemp_change	Percentage point change in unemployment (unemp_lag1 – unemp_lag2)
real_gdp_growth	Growth of real GPD, percent change from previous year (based on CPDS)
real_gdp_growth_lag1	real_gdp_growth, lagged by one year
real_gdp_growth_lag2	real_gdp_growth, lagged by two years
type	Type of manifesto (annotated vs not annotated)
cabinet_status	Cabinet status AFTER election t
cabinet_status_lag	Cabinet status AT TIME WHEN election manifesto was published
populism_popu_list	Populist party, according to PopuList project
populist_dummy	Dummy variable for populist vs non-populist parrties
populism_popu_list_categories	Populist party, fine-grained PopuList coding
vote_share_cmp	Vote share at election t
vote_share_cmp_lag	Vote share at election t-1
nostalgia_sentences_per_1000_svm	Nostalgia: SVM
nostalgia_sentences_per_1000	Nostalgic: base dictionary
nostalgia_sentences_per_1000_bert	Nostalgia: DistilBERT
nostalgia_sentences_per_1000_emb	Nostalgia: base dictionary + embeddings
nostalgia_sentences_per_1000_sentiment	Nostalgia: base dictionary + sentiment
nostalgia_sentences_per_1000_sentiment_emb	Nostalgia: base dictionary + embeddings + sentiment
nost_sents_sum_base	Sum of nostalgic sentences: base dictionary
nost_sents_sum_emb	Sum of nostalgic sentences: base dictionary + embeddings
nost_sents_sum_svm	Sum of nostalgic sentences: SVM
nost_sents_sum_bert	Sum of nostalgic sentences: DistilBERT
nost_sents_sum_emb_pos	Sum of nostalgic sentences: base dictionary + embeddings + entiment
nost_sents_sum_pos	Sum of nostalgic sentences: base dictionary + sentiment
n_sentences_manifesto	Number of manifesto sentences (limited to sentences from 6-59 words)
rile	Left-right position (RILE)
logrile	Logged RILE (see Lowe et al. 2011)
stateconomy	State Involvement in the Economy (see Lowe et al. 2011)
loglibcons_base	Cultural conservatism (based on Lowe et al.'s (2011) original scale)
loglibcons	Cultural conservatism (based extended scale)
decade	Decade
nominal_gdp_growth	Growth of nominal GPD, percent change from previous year (based on CPDS)
nominal_gdp_growth_lag1	nominal_gdp_growth, lagged by one year
nominal_gdp_growth_lag2	nominal_gdp_growth, lagged by two years
cabinet_status_lag2	Binary classification of cabinet_status_lag into Government and Opposition

3 Sentence-Level Datasets

In addition to the manifesto-level dataset, we provide various datasets at the (quasi-)sentence level, which we describe below. These datasets contain the (translated) text and metadata for each observation.

PolNos Version 1.0 translates manifestos included in `Version 2019-2` of the Manifesto Corpus (Krause et al. 2019). If you use one or more sentence-level datasets, please cite Krause et al. (2019). If you retrieve data through the *manifestoR* package, we strongly recommend citing Lewandowski, Merz, and Regel (2020).

3.1 *PolNos* Sentence-Level Data

The dataset `data_polnos_sentencelevel.rds` contains the machine-translated party manifestos along with the coding of nostalgia according to the six measures. The translations are on the level of segmented sentences (using the *quanteda* R package) or based on the existing quasi-sentence coding by the Manifesto Project (Merz, Regel, and Lewandowski 2016). Since we focus on the level of sentences, we excluded very short and long "sentences" and limit the dataset to sentences ranging from 6 to 59 words.[1] These may either be quasi-sentences containing few words, incorrectly segmented sentences, or overly long sentences. Table 2 describes all variables included in the sentence-level dataset.

In this section, we show how to load the dataset into R and describe the variables in this dataset. The data frame `data_polnos_sentencelevel` contains 1,192,680 observations. Each sentence is a quasi-sentence or a natural sentence. 708,463 sentences contain CMP policy codes. The remaining 484,217 sentences were split into natural sentences based on punctuation characters and do not include information on the policy area.

3.2 Hand-Coded Manifesto Sentences

The *PolNos* datasets also includes the random sample of 1,200 manifesto sentences, which were annotated by four coders. We selected a random sample of 25 sentences per country that were classified as nostalgic rhetoric according to the broadest method (Dictionary + Embeddings) and 25 sentences per country that were not coded as nostalgic according to this dictionary. After two initial reliability exercises of 100 sentences each, four instructed coders annotated each of the 1,200 sentences. More details, inter-coder reliability statistics, and the coding instructions are included in the accompanying paper. The dataset is called `data_polnos_handcoding.csv`. Table 4 describes all variables included in this dataset.

[1] Punctuation characters are not counted as words.

Table 2: Variable description: Sentence-level PolNos dataset

Variable	Details
manifesto_id	ID for each manifesto (party ID + election date)
type	Type of manifesto (annotated vs not annotated)
text	Machine-translated text (translated using Google Translate)
doc_id	Document-identifier within each manifesto
countryname	Country name
party	Party identifier (from Manifesto Project)
edate	Date of election (format: YYYY-MM-DD)
pos	Position variable from the Manifesto Corpus for quasi-sentence annotated manifestos
cmp_code	CMP code of quasi-sentence (if type == 'Annotated')
ntoken	Number of tokens in sentence (excluding punctuation characters)
sentiment	Sentiment of sentences (Proksch et al. (2019) aggregation
nostalgia_sentence_dummy	Nostalgic according to Base Dictionary
nostalgia_sentence_dummy_emb	Nostalgic according to Base Dictionary + Embeddings
nostalgia_sentence_dummy_sentiment	Nostalgic according to Base Dictionary + Positive Sentiment
nostalgia_sentence_dummy_sentiment_emb	Nostalgic according to Base Dictionary + Embeddings + Positive Sentiment
nostalgia_sentence_bert	Nostalgic according to DistilBERT classifier
nostalgia_sentence_svm	Nostalgic according to SVM classifier
nostalgia_sentence_svm_or_bert	Nostalgic according to SVM OR DistilBERT
nostalgia_sentence_ml_and_emb	Nostalgic according to SVM OR DistilBERT AND the broadest dictionary
nostalgia_ensemble_count	Count of nostalgic classifications (out of 6)

3.3 Hand-Coded Manifesto Sentences for Comparison of Measures

We also include 3,515 sentences, which have been classified as nostalgic by one or more of the six methods. More details on the sampling method are provided in SI Section D of the main paper. All of these sentences have subsequently been coded by two human coders. Moreover, we used GPT 3.5 to classify each sentence as nostalgic (on a scale from 1 to 10 and by providing a justification for the score) after providing minimal instructions. The dataset is called `data_polnos_handcoding_validation.csv`. Table 2 describes all variables included in the sentence-level dataset.

3.4 Full Machine-Translated Corpus

We also provide the dataset `data_manifestostranslated.rds` which does not filter sentences based on text length. This dataset does not contain the codings of nostalgia, but could be used by other researchers who want to work with the full set of translated sentences. It includes very short sentences and very long sentences. We strongly advise researchers to assess the raw translations very thoroughly before conducting analyses. Table 5 describes all variables included in the sentence-level dataset.

The data frame `data_polnos_translated_full` contains 1,439,551 observations. Each sentence is a quasi-sentence or a natural sentence. We limit the dataset to sentences ranging from 6 to 59 words. 805,245 sentences contain CMP policy codes. The remaining 634,306 sentences were split into natural sentences based on punctuation characters and do not include information on the policy area.

Table 3: Variable description: Dataset of coded manifesto sentences

Variable	Details
doc_id	Document-identifier within each manifesto
countryname	Country name
party	Party identifier (from Manifesto Project)
manifesto_id	ID for each manifesto (party ID + election date)
text	Manifesto sentence (translated for non-English manifestos)
cmp_code	CMP code from Manifesto Project dataset
nostalgic_at_least_1	Nostaglic sentence (according to at least 1 coder)
nostalgic_at_least_2	Nostaglic sentence (according to at least 2 coders)
nostalgic_at_least_3	Nostaglic sentence (according to at least 3 coders)
nostalgic_at_least_4	Nostaglic sentence (according to all 4 coders)
translation_at_least_1	Incomprehensible translation (according to at least 1 coder)
translation_at_least_2	Incomprehensible translation (according to at least 2 coders)
translation_at_least_3	Incomprehensible translation (according to at least 3 coders)
translation_at_least_4	Incomprehensible translation (according to all 4 coders)
nostalgic_coder1	Nostalgic rhetoric according to Coder 1
nostalgic_coder2	Nostalgic rhetoric according to Coder 2
nostalgic_coder3	Nostalgic rhetoric according to Coder 3
nostalgic_coder4	Nostalgic rhetoric according to Coder 4
translation_coder1	Incomprehensible translation according to Coder 1
translation_coder2	Incomprehensible translation according to Coder 2
translation_coder3	Incomprehensible translation according to Coder 3
translation_coder4	Incomprehensible translation according to Coder 4
translation_agreement_coders	Number of coders who labeled translation as incomprehensible
nostalgia_sum	Count of nostalgic terms (base dictionary)
nostalgia_sum_emb	Count of nostalgic terms (base dictionary + embeddings)
nostalgia_emb	Binary indicator of nostalgia (embeddings-based dictionary)
nostalgia_dictionary_emb	Binary indicator of nostalgia (base dictionary + embeddings)
nostalgia_dictionary	Binary indicator of nostalgia (base dictionary)
nostalgia_agreement_coders	Number of coders who labeled sentence as nostalgic rhetoric

4 Nostalgia Dictionary

The repository contains the nostalgia dictionary in three formats:

- A **quanteda** dictionary object: `data_dictionary_nostalgia.rds`.
- A data frame in a spreadsheet format, stored as a UTF-8 encoded `csv` file: `data_dictionary_nostalgia.csv`
- A YAML file, which can be edited in text editors and imported into **quanteda** again: `data_dictionary_nostalgia.yml`

The dictionary consists of two overarching keys: the key `nostalgia` contains the words and phrases from the "base" dictionary. This key is divided up into features from Davalos et al. (2015), extensions and adjustments for political texts (`nostalgia_extensions`), and hand-coding of sentences `nostalgia_handcoding` at the initial stage of the project. The second key, `nostalgia_emb`, contains the terms selected through manually coding the nearest neighbours – based on word embeddings – (Hvitfeldt and Silge 2021: ch. 5) of the terms from words and phrases included in three `nostalgia` sub-keys Details on the dictionary creation are provided in the Online Supporting Information (Section B) of Müller and Proksch (*British Journal of Political Science*). Below, we show how to load and apply the dictionary.

Table 4: Variable description: Hand-coded manifesto sentences for final validation of measures

Variable	Details
manifesto_id	ID for each manifesto (party ID + election date)
countryname	Country name
text_pre	Sentence preceeding sentence stored in 'text' column
text	Machine-translated text (translated using Google Translate)
text_post	Sentence succeeding sentence stored in 'text' column
party	Party identifier (from Manifesto Project)
party_family_recoded	Party family (according to manifesto project)
cmp_code	CMP code
nostalgia_sentence_dummy_emb	Nostalgic according to Base + Embeddings Dictionary
nostalgia_sentence_bert	Nostalgic according to DistilBERT classifier
nostalgia_sentence_svm	Nostalgic according to SVM classifier
nostalgic_coding_coder1	Nostalgic sentence according to Coder 1
nostalgic_coding_coder2	Nostalgic sentence according to Coder 2
nostalgia_coded_both	Nostalgic sentence according to both coders
nostalgia_coded_at_least_one	Nostalgic sentence according to at least one coder
score_gpt	Nostalgia score based on GPT 3.5 rating (ranging from 1 to 10)
justification_gpt	Justification for rating

Table 5: Variable description: Full machine-translated text corpus

Variable	Details
manifesto_id	ID for each manifesto (party ID + election date)
type	Type of manifesto (annotated vs not annotated)
text	Machine-translated text (translated using Google Translate)
doc_id	Document-identifier within each manifesto
countryname	Country name
party	Party identifier (from Manifesto Project)
edate	Date of election (format: YYYY-MM-DD)
pos	Position variable from the Manifesto Corpus for quasi-sentence annotated manifestos
cmp_code	CMP code assigned by Manifesto Project coders

```r
# load the quanteda package and the dictionary
library(quanteda)
data_dictionary_nostalgia <- readRDS("data_dictionary_nostalgia.rds")

# inspect structure of dictionary object
data_dictionary_nostalgia
```

```
Dictionary object with 2 primary key entries and 2 nested levels.
- [nostalgia]:
  - [nostalgia_davalos]:
    - remember* when, memory lane, flashback, go back in time, good old day*, miss those days
  - [nostalgia_handcoding]:
    - heritage, historic, history, comfortable with * past, for * of years, for * of decades
  - [nostalgia_extensions]:
    - unforgettable, memory, memorable, good time*, time ago, not forget, past, year* ago, pi
- [nostalgia_emb]:
```

- commemoration, conserve, continued, continuous, historical, legacy, memorial, memories, c

```
# application: apply dictionary to a corpus object
corp <- corpus(c(
    doc1 = "Let's remember the good times",
    doc2 = "We need to preserve our cultural heritage",
    doc3 = "An irrelevant document"))


# get dictionary counts on the level of documents
# and convert to data frame
corp |>
    tokens() |>
    tokens_lookup(dictionary = data_dictionary_nostalgia,
                  levels = 1) |>
    dfm() |>
    convert(to = "data.frame")
```

```
  doc_id nostalgia nostalgia_emb
1   doc1         1             0
2   doc2         2             0
3   doc3         0             0
```

```
# show which words or phrases have been identified through the dictionary
corp |>
    tokens() |>
    tokens_lookup(dictionary = data_dictionary_nostalgia,
                  levels = 1, exclusive = FALSE)
```

```
Tokens consisting of 3 documents.
doc1 :
[1] "Let's"     "remember"  "the"       "NOSTALGIA"

doc2 :
[1] "We"        "need"      "to"        "NOSTALGIA" "our"       "cultural"
[7] "NOSTALGIA"

doc3 :
[1] "An"         "irrelevant" "document"
```

4.1 PolNos Classifiers

The fine-tuned DistilBERT classifier and a tutorial on how to predict sentences with this classifier and store them as a csv file are provided in the replication materials of Müller and Proksch (BJPS):

> Müller, Stefan and Sven-Oliver Proksch. 2023. *Replication Data for: Nostalgia in European Party Politics: A Text-Based Measurement Approach.* Harvard Dataverse V1. DOI: 10.7910/DVN/W8VGJF

`0000_README.pdf` in the replication materials of Müller and Proksch (BJPS) contains details on all files.\ The script `07_tutorial_classify_nostalgia_distilbert.ipynb` provides a tutorial on how to apply the fine-tuned DistilBERT model to new data and save the sentence-level predictions.

5 Load *PolNos* Datasets

In this section, we show how to load *PolNos* datasets. For these examples, we use the R statistical programming language.

```
# load manifesto-level dataset
data_polnos_manifestolevel <- readRDS("data_polnos_manifestolevel.rds")

# load sentence-level dataset
data_polnos_sentencelevel <- readRDS("data_polnos_sentencelevel.rds")

# load full machine-translated dataset
data_translated_full <- readRDS("data_translated_full.rds")

# load hand-coded sample of manifesto sentence
data_polnos_handcoding <- read.csv("data_polnos_handcoding.csv",
                                   fileEncoding = "utf-8")
```

6 Merge *PolNos* with *Manifesto Project Dataset*

We can easily merge *PolNos* with the full *Manifesto Project Dataset*. In this example, we extract the party ID, election date, and measures of nostalgic rhetoric, and merge this subsetted dataset with the *Manifesto Project Dataset* (`Version 2022a`).

6.1 Load and Prepare *Manifesto Project Dataset*

```r
# load dplyr package for data wrangling
library(dplyr)

# load Manifesto Project dataset
# register and download at https://manifesto-project.wzb.eu/datasets
data_manifestoproject <- read.csv("MPDataset_MPDS2022a.csv")

# recode party as character
data_manifestoproject$party <- as.character(data_manifestoproject$party)

# transform date format to YYYY-MM-DD
data_manifestoproject$edate <- as.Date(data_manifestoproject$edate, "%d/%m/%Y")

# recode date as character vector for merging purposes
data_manifestoproject$edate <- as.character(data_manifestoproject$edate)
```

6.2 Merge Manifesto-level *PolNos* Dataset with *Manifesto Project Dataset*

```r
# select party and edate for merging and all variables starting with "nost"
data_polnos_manifestolevel_select <- data_polnos_manifestolevel |>
    dplyr::select(party, edate, starts_with("nost"))

# recode party variable as character
data_polnos_manifestolevel_select$party <-
    as.character(data_polnos_manifestolevel_select$party)

names(data_polnos_manifestolevel_select)
```

```
 [1] "party"
 [2] "edate"
 [3] "nostalgia_main"
 [4] "nostalgia_sentences_per_1000_svm"
 [5] "nostalgia_sentences_per_1000"
 [6] "nostalgia_sentences_per_1000_bert"
 [7] "nostalgia_sentences_per_1000_emb"
 [8] "nostalgia_sentences_per_1000_sentiment"
 [9] "nostalgia_sentences_per_1000_sentiment_emb"
[10] "nost_sents_sum_base"
```

```
[11] "nost_sents_sum_emb"
[12] "nost_sents_sum_svm"
[13] "nost_sents_sum_bert"
[14] "nost_sents_sum_emb_pos"
[15] "nost_sents_sum_pos"
```

```
data_polnos_manifestolevel_merged <- data_polnos_manifestolevel_select |>
    dplyr::left_join(data_manifestoproject, by = c("party", "edate"))
```

The number of variables increases from 15 to 187. The number of observations (=manifestos) remains the same.

6.3 Merge Sentence-level *PolNos* Dataset with *Manifesto Project Dataset*

```
# select party and edate for merging and all variables starting with "nost*"
data_polnos_sentencelevel_select <- data_polnos_sentencelevel |>
    dplyr::select(party, edate, starts_with("nost"))

names(data_polnos_sentencelevel_select)
```

```
 [1] "party"
 [2] "edate"
 [3] "nostalgia_sentence_dummy"
 [4] "nostalgia_sentence_dummy_emb"
 [5] "nostalgia_sentence_dummy_sentiment"
 [6] "nostalgia_sentence_dummy_sentiment_emb"
 [7] "nostalgia_sentence_bert"
 [8] "nostalgia_sentence_svm"
 [9] "nostalgia_sentence_svm_or_bert"
[10] "nostalgia_sentence_ml_and_emb"
[11] "nostalgia_ensemble_count"
```

```
# recode party variable as character
data_polnos_sentencelevel_select$party <-
    as.character(data_polnos_sentencelevel_select$party)

data_polnos_sentencelevel_merged <- data_polnos_sentencelevel_select |>
    dplyr::left_join(data_manifestoproject, by = c("party", "edate"))
```

```r
# compare rows and columns in both datasets
dim(data_polnos_sentencelevel_select)
```

```
[1] 1192680      11
```

```r
dim(data_polnos_sentencelevel_merged)
```

```
[1] 1192680     183
```

The number of variables increases from 11 to 183. As expected, the number of observations (=sentences) remains unchanged.

References

Armingeon, Klaus, Sarah Engler, and Lucas Leemann. 2022. *Comparative Political Data Set 1960–2020*. https://www.cpds-data.org.

Davalos, Sergio, Altaf Merchant, Gregory M. Rose, Brenton J. Lessley, and Ankur M. Teredesai. 2015. "'The Good Old Days': An Examination of Nostalgia in Facebook Posts." *International Journal of Human-Computer Studies* 83: 83–93. https://doi.org/10.1016/j.ijhcs.2015.05.009.

Döring, Holger, and Philip Manow. 2021. *Parliaments and Governments Database (ParlGov): Information on Parties, Elections and Cabinets in Modern Democracies. Development Version*. http://parlgov.org.

Hvitfeldt, Emil, and Julia Silge. 2021. *Supervised Machine Learning for Text Analysis in R*. Boca Raton: CRC Press. https://smltar.com.

Krause, Werner, Pola Lehmann, Jirka Lewandowski, Theres Matthieß, Nicolas Merz, Sven Regel, and Annika Werner. 2019. *Manifesto Corpus: Version 2019-2*. Berlin: WZB Berlin Social Science Center. https://manifesto-project.wzb.eu/information/documents/corpus.

Lewandowski, Jirka, Nicolas Merz, and Sven Regel. 2020. *manifestoR: Access and Process Data and Documents of the Manifesto Project*. https://CRAN.R-project.org/package=manifestoR.

Merz, Nicolas, Sven Regel, and Jirka Lewandowski. 2016. "The Manifesto Corpus: A New Resource for Research on Political Parties and Quantitative Text Analysis." *Research & Politics* 3 (2): 1–8. https://doi.org/10.1177/2053168016643346.

Volkens, Andrea, Tobias Burst, Werner Krause, Pola Lehmann, Theres Matthieß, Nicolas Merz, Sven Regel, Bernhard Weßels, and Lisa Zehnter. 2021. *The Manifesto Data Collection: Manifesto Project (MRG/CMP/MARPOR). Version 2021a*. https://doi.org/10.25522/manifesto.mpds.2021a.