PImPo - Parties' Immigration and Integration Positions Dataset

Data Documentation and Codebook

Pola Lehmann & Malisa Zobel

January 2018

The PImPo dataset is released with doi.org/10.25522/manifesto.pimpo.2018. When using the dataset please cite Pola Lehmann, Malisa Zobel (2018): Positions and Saliency of Immigration in Party Manifestos. A Novel Dataset using Crowd Coding. European Journal for Political Research. Online First 2018. doi.org/10.1111/1475-6765.12266 If you use the dataset on the quasi sentence level, please additionally cite Lehmann, Pola, Theres Matthieß, Nicolas Merz, Sven Regel, Annika Werner (2015): Manifesto Corpus. Version: 20150708174629. Berlin: WZB Berlin Social Science Center.

Content

General information

The Parties' Immigration and Integration Positions Dataset (PImPo) includes data on parties' immigration and integration positions and saliency in 14 countries (Australia, Austria, Canada, Sweden, Denmark, Finland, Germany, Ireland, Netherlands, New Zealand, Norway, Spain, Switzerland, USA) between 1998 and 2013 based on crowd coding of parties' election manifestos. The dataset entails variables on immigration and integration positions and saliency for each party in one election.

After a simple merge with the original verbatim from the Manifesto Corpus (Lehmann, Pola, Theres Matthieß, Nicolas Merz, Sven Regel, Annika Werner (2015): Manifesto Corpus. Version: 20150708174629. Berlin: WZB Berlin Social Science Center.) it also gives data on the quasi-sentence level. Variables here include information on the topic of the quasi-sentence (immigration, integration, or none of both), the direction (supportive, sceptical, neutral), as well as information on agreement levels between crowd coders for each quasi-sentence.

Variables are named the same no matter in which format you open the data (xlsx;dta; csv) but the variables are only labeled in the Stata format (.dta). You can however use the following codebook to know which variables contain what information. It, however, gives only a succinct overview, for detailed information on the dataset and its generation please refer to the accompanying article:

Pola Lehmann, Malisa Zobel (2018): Positions and Saliency of Immigration in Party Manifestos. A Novel Dataset using Crowd Coding. European Journal for Political Research. Online First.

1. Data structure

Data is available either with the party as the unit of observation or with the quasi-sentence as the unit of observation.

1.1. PImPo – Party level dataset

The unit of observation is the party. Party and date uniquely identify each observation.

Detailed information on each variable

Country

```
-------------------------------------------------------------------------------
country                                                   MARPOR country id
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 label:  country
                 range:  [11,64]                      units:  1
          unique values:  14                      missing .:  0/242
           value labels:  11  Sweden
                          12  Norway
                          13  Denmark
                          14. Finland
                          22. Netherlands
                          33. Spain
                          41. Germany
                          42. Austria
                          43. Switzerland
```

```
                        53. Ireland
                        61. USA
                        62. Canada
                        63. Australia
                        64. New Zealand
```

Comment: the country codes are the same as in the MARPOR datasets.

Party

```
-------------------------------------------------------------------------------
party                                                          MARPOR party id
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 label:  parties
                 range:  [11110,64902]                units:  1
          unique values:  110                    missing .:  0/242
           value labels:
                       11110 swe: Green Ecology Party
                       11320 swe: Social Democratic Labour Party
                       11420 swe: Liberal People's Party
                       11220 swe: Left Communists Party
                       11520 swe: Christian Democratic Community Party
                       11620 swe: Moderate Coalition Party
                       11710 swe: Sweden Democrats
                       11810 swe: Centre Party
                       12221 nor: Socialist Left Party
                       12320 nor: Norwegian Labour Party
                       12420 nor: Liberal Party
                       12520 nor: Christian People's Party
                       12620 nor: Conservative Party
                       12810 nor: Centre Party
                       12951 nor: Progress Party
                       13001 dnk: Liberal Alliance
                       13229 dnk: Red-Green Unity List
                       13230 dnk: Socialist People's Party
                       13320 dnk: Social Democratic Party
                       13330 dnk: Centre Democrats
                       13410 dnk: Radical Party
                       13420 dnk: Liberals
                       13520 dnk: Christian People's Party
                       13620 dnk: Conservative People's Party
                       13720 dnk: Danish People's Party
                       13951 dnk: Progress Party
                       14110 fin: Green Union
                       14223 fin: Left Wing Alliance
                       14320 fin: Finnish Social Democrats
                       14520 fin: Christian Democrats in Finland
                       14620 fin: National Coalition
                       14810 fin: Finnish Centre
                       14820 fin: Finnish Rural Party
                       14901 fin: Swedish People's Party
                       22110 nld: Green Left
                       22220 nld: Socialist Party
                       22320 nld: Labour Party
                       22330 nld: Democrats'66
                       22420 nld: People's Party for Freedom and Democracy
                       22521 nld: Christian Democratic Appeal
                       22526 nld: Christian Union
                       22722 nld: Party of Freedom
                       22951 nld: Party for the Animals
                       22952 nld: Reformed Political Party
                       33091 esp: Future Yes
```

33092 esp: Amaiur
33093 esp: Commitment-Q
33220 esp: United Left
33320 esp: Spanish Socialist Workers' Party
33440 esp: Union, Progress and Democracy
33610 esp: Popular Party
33611 esp: Convergence and Union
33612 esp: Forum Asturias
33902 esp: Basque Nationalist Party
33903 esp: Basque Solidarity
33907 esp: Canarian Coalition
33909 esp: Aragonist Council
41113 deu: Alliance'90/Greens
41221 deu: Party of Democratic Socialism
41223 deu: The Left
41320 deu: Social Democratic Party of Germany
41420 deu: Free Democratic Party
41521 deu: Christian Democratic Union/C.Social Union
42110 aut: The Greens
42110 aut: The Greens
42110 aut: The Greens
42110 aut: The Greens
42220 aut: Austrian Communist Party
42320 aut: Austrian Social Democratic Party
42320 aut: Austrian Social Democratic Party
42320 aut: Austrian Social Democratic Party
42320 aut: Austrian Social Democratic Party
42420 aut: Austrian Freedom Party
42420 aut: Austrian Freedom Party
42420 aut: Austrian Freedom Party
42420 aut: Austrian Freedom Party
42520 aut: Austrian People's Party
42520 aut: Austrian People's Party
42520 aut: Austrian People's Party
42520 aut: Austrian People's Party
42710 aut: Alliance for the Future of Austria
42710 aut: Alliance for the Future of Austria
43110 che: Green Party of Switzerland
43120 che: Green Liberal Party
43220 che: Swiss Labour Party
43320 che: Social Democratic Party of Switzerland
43420 che: Radical Democratic Party
43520 che: Conservative Christian Social Party
43530 che: Protestant People's Party of Switzerland
43540 che: Christian Social Party
43711 che: Federal Democratic Union
43810 che: Swiss People's Party
43811 che: Conservative Democratic Party of Switzer.
53021 irl: United Left Alliance
53110 irl: Green Party
53230 irl: Socialist Party
53320 irl: Labour Party
53420 irl: Progressive Democrats
53520 irl: Familiy of the Irish
53620 irl: Soldiers of Destiny
53951 irl: Ourselves Alone
61320 usa: Democratic Party
61620 usa: Republican Party
62110 can: Green Party
62320 can: New Democratic Party
62420 can: Liberal Party of Canada
62623 can: Conservative Party of Canada

```
                    62901 can: Quebec Bloc
                    63110 aus: Australian Greens
                    63320 aus: Australian Labor Party
                    63410 aus: Palmer United Party
                    63620 aus: Liberal Party of Australia
                    63710 aus: Katter's Australian Party
                    63810 aus: National Country Party
                    64110 nzl: Green Party of Aotearoa New Zealand
                    64320 nzl: New Zealand Labour Party
                    64420 nzl: ACT New Zealand
                    64421 nzl: United Future New Zealand
                    64422 nzl: Progressive Party
                    64620 nzl: New Zealand National Party
                    64621 nzl: New Zealand First Party
                    64901 nzl: Maori Party
                    64902 nzl: Mana Party
```

Comment: the party codes are the same as in the MARPOR datasets.

Date

```
-------------------------------------------------------------------------------
date                                                            Election date
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 range:  [199803,201309]              units:  1
         unique values:  32                    missing .:  0/242
                  mean:    200741
             std. dev:    390.713
           percentiles:        10%       25%       50%       75%       90%
                            200209    200609    200803    201104    201111
```

Totals

```
-------------------------------------------------------------------------------
totals                                                        Total QS coded
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 range:  [20,6330]                    units:  1
         unique values:  221                   missing .:  0/242
                  mean:    972.533
             std. dev:    1021.69
           percentiles:        10%       25%       50%       75%       90%
                                88       204       594      1576      2225
```

Totals_immi

```
-------------------------------------------------------------------------------
totals_immi                                     Total QS on immigration
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 range:  [1,119]                      units:  1
         unique values:  57                    missing .:  42/242
                  mean:    19.495
             std. dev:    19.3292
           percentiles:        10%       25%       50%       75%       90%
                                 2         5        13        29      43.5
```

Totals_inti

```
-------------------------------------------------------------------------------
totals_inti                                     Total QS on integration
-------------------------------------------------------------------------------
```

```
                type:  numeric (long)
               range:  [1,169]                      units:  1
        unique values:  67                       missing .:  34/242
                mean:   24.3558
            std. dev:   26.9209
          percentiles:        10%       25%       50%       75%       90%
                               1         5        15        37        62
```

Saliency

```
-------------------------------------------------------------------------------
saliency                                    Overall saliency immi + inte
-------------------------------------------------------------------------------
                type:  numeric (double)
               range:  [0,33.333333]                units:  1.000e-10
        unique values:  215                      missing .:  0/242
                mean:    4.7195
            std. dev:   5.39263
          percentiles:        10%       25%       50%       75%       90%
                               0   1.31752    3.4474   5.65111   10.2767
```

Comments: saliency is calculated as the proportion of immigration and integration related quasi-sentences to the total number of quasi-sentences.

Saliency_immi

```
-------------------------------------------------------------------------------
saliency_immi                                        Saliency immigration
-------------------------------------------------------------------------------
                type:  numeric (double)
               range:  [0,27.5]                     units:  1.000e-09
        unique values:  196                      missing .:  0/242
                mean:   2.29299
            std. dev:   3.47259
          percentiles:        10%       25%       50%       75%       90%
                               0   .466667   1.24018   2.51366   5.19481
```

Comments: saliency of immigration is calculated as the proportion of immigration related quasi-sentences to the total number of quasi-sentences.

Saliency_inti

```
-------------------------------------------------------------------------------
saliency_inti                                        Saliency integration
-------------------------------------------------------------------------------
                type:  numeric (double)
               range:  [0,18.604651]                units:  1.000e-09
        unique values:  207                      missing .:  0/242
                mean:   2.42651
            std. dev:   2.65673
          percentiles:        10%       25%       50%       75%       90%
                               0   .580271   1.82504   3.27103   5.20231
```

Comments: saliency of integration is calculated as the proportion of integration related quasi-sentences to the total number of quasi-sentences.

Immi_pos

```
-------------------------------------------------------------------------------
immi_pos                                             Immigration position
-------------------------------------------------------------------------------
                type:  numeric (double)
               range:  [-1,1]                       units:  1.000e-09
```

```
        unique values:  115                      missing .:  42/242
                 mean:   .205991
            std. dev:    .59069
          percentiles:        10%       25%       50%       75%       90%
                           -.69697  -.184615   .307692   .714286         1
```

Comments: immigration position is calculated by subtracting the share of sceptical quasi-sentences from the share of supportive quasi-sentences and dividing this by the share of these two plus the share of the neutral quasi-sentences.

$$immi\_pos = \frac{\sum_{p,i} QS_{immi+} - \sum_{p,i} QS_{immi-}}{\sum_{p,i} QS_{immi+} + \sum_{p,i} QS_{immi-} \sum_{p,i} QS_{immi\_neu}}$$

p = party, i = election, immi+ = all positive immigration related quasi-sentences, immi- = all negative immigration related quasi-sentences, immi_neu = all neutral/status quo immigration related quasi-sentences

Immi_pos_saliency

```
-------------------------------------------------------------------------------
immi_pos_saliency                                                 (unlabeled)
-------------------------------------------------------------------------------
                 type:  numeric (double)
                range:  [0,100]                      units:  1.000e-07
        unique values:  79                       missing .:  42/242
                 mean:   83.4422
            std. dev:   18.1109
          percentiles:        10%       25%       50%       75%       90%
                           60.8187        75      87.5       100       100
```

Comments: gives the percentage of directional quasi-sentences. It is calculated by adding the supportive and the sceptical immigration quasi-sentences and dividing them by all immigration quasi-sentence (supportive, sceptical and neutral).

Inti_pos

```
-------------------------------------------------------------------------------
inti_pos                                                Integration position
-------------------------------------------------------------------------------
                 type:  numeric (double)
                range:  [-1,1]                       units:  1.000e-09
        unique values:  110                      missing .:  34/242
                 mean:   .566126
            std. dev:   .534105
          percentiles:        10%       25%       50%       75%       90%
                          -.152174   .355994   .761387   .965689         1
```

Comments: integration position is calculated by subtracting the share of sceptical quasi-sentences from the share of supportive quasi-sentences and dividing this by the share of these two plus the share of the neutral quasi-sentences.

$$immi\_pos = \frac{\sum_{p,i} QS_{inti+} - \sum_{p,i} QS_{inti-}}{\sum_{p,i} QS_{inti+} + \sum_{p,i} QS_{inti-} \sum_{p,i} QS_{inti\_neu}}$$

p = party, i = election, inti+ = all positive integration related quasi-sentences, inti- = all negative integration related quasi-sentences, inti_neu = all neutral/status quo integration related quasi-sentences

Inti_pos_saliency

```
-------------------------------------------------------------------------------
inti_pos_saliency                                                 (unlabeled)
-------------------------------------------------------------------------------
                  type:  numeric (double)
                 range:  [0,100]                      units:  1.000e-07
          unique values:  88                     missing .:  34/242
                  mean:   90.2323
             std. dev:   15.3232
           percentiles:        10%       25%       50%       75%       90%
                            76.4706   86.5766   95.2381       100       100
```

Comments: gives the percentage of directional quasi-sentences. It is calculated by adding the supportive and the sceptical integration quasi-sentences and dividing them by all integration quasi-sentence (supportive, sceptical and neutral)

1.2. PImPo – Quasi-sentence level dataset

The crowd coding was done on the quasi-sentence level. Quasi-sentences are units of analysis capturing one unique argument in a sentence. They are never longer than one natural sentence, but sometimes shorter than a natural sentence. For more information see Werner, Annika, Andrea Volkens, Onawa Lacewell (2015): Manifesto Coding Instructions. 5th revised edition. Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB).

The Quasi-sentence level dataset gives information on how the crowd coded each individual quasi-sentence. For copyright reasons it does not contain the original verbatim of the quasi-sentence, but the verbatim can be very easily merged to the dataset from the Manifesto Corpus. To make the merge as easy as possible, we provide a R-Script which adds the verbatim. To execute this R-Script no prior R-knowledge is required.

Country

```
-------------------------------------------------------------------------------
country                                                 MARPOR country id
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 label:  country
                 range:  [11,64]                       units:  1
          unique values:  14                       missing: 0/235353

           value labels:  11  Sweden
                          12  Norway
                          13  Denmark
                          14. Finland
                          22. Netherlands
                          33. Spain
                          41. Germany
                          42. Austria
                          43. Switzerland
                          53. Ireland
                          61. USA
                          62. Canada
                          63. Australia
                          64. New Zealand
```

Comment: the country codes are the same as in the MARPOR datasets.

Party

```
-------------------------------------------------------------------------------
party                                                          MARPOR party id
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 label:  parties
                 range:  [11110,64902]                units:  1
         unique values:  110                      missing .: 0/235353
          value labels:
                        11110 swe: Green Ecology Party
                        11320 swe: Social Democratic Labour Party
                        11420 swe: Liberal People's Party
                        11220 swe: Left Communists Party
                        11520 swe: Christian Democratic Community Party
                        11620 swe: Moderate Coalition Party
                        11710 swe: Sweden Democrats
                        11810 swe: Centre Party
                        12221 nor: Socialist Left Party
                        12320 nor: Norwegian Labour Party
                        12420 nor: Liberal Party
                        12520 nor: Christian People's Party
                        12620 nor: Conservative Party
                        12810 nor: Centre Party
                        12951 nor: Progress Party
                        13001 dnk: Liberal Alliance
                        13229 dnk: Red-Green Unity List
                        13230 dnk: Socialist People's Party
                        13320 dnk: Social Democratic Party
                        13330 dnk: Centre Democrats
                        13410 dnk: Radical Party
                        13420 dnk: Liberals
                        13520 dnk: Christian People's Party
                        13620 dnk: Conservative People's Party
                        13720 dnk: Danish People's Party
                        13951 dnk: Progress Party
                        14110 fin: Green Union
                        14223 fin: Left Wing Alliance
                        14320 fin: Finnish Social Democrats
                        14520 fin: Christian Democrats in Finland
                        14620 fin: National Coalition
                        14810 fin: Finnish Centre
                        14820 fin: Finnish Rural Party
                        14901 fin: Swedish People's Party
                        22110 nld: Green Left
                        22220 nld: Socialist Party
                        22320 nld: Labour Party
                        22330 nld: Democrats'66
                        22420 nld: People's Party for Freedom and Democracy
                        22521 nld: Christian Democratic Appeal
                        22526 nld: Christian Union
                        22722 nld: Party of Freedom
                        22951 nld: Party for the Animals
                        22952 nld: Reformed Political Party
                        33091 esp: Future Yes
                        33092 esp: Amaiur
                        33093 esp: Commitment-Q
                        33220 esp: United Left
                        33320 esp: Spanish Socialist Workers' Party
                        33440 esp: Union, Progress and Democracy
```

33610 esp: Popular Party
33611 esp: Convergence and Union
33612 esp: Forum Asturias
33902 esp: Basque Nationalist Party
33903 esp: Basque Solidarity
33907 esp: Canarian Coalition
33909 esp: Aragonist Council
41113 deu: Alliance‘90/Greens
41221 deu: Party of Democratic Socialism
41223 deu: The Left
41320 deu: Social Democratic Party of Germany
41420 deu: Free Democratic Party
41521 deu: Christian Democratic Union/C.Social Union
42110 aut: The Greens
42110 aut: The Greens
42110 aut: The Greens
42110 aut: The Greens
42220 aut: Austrian Communist Party
42320 aut: Austrian Social Democratic Party
42320 aut: Austrian Social Democratic Party
42320 aut: Austrian Social Democratic Party
42320 aut: Austrian Social Democratic Party
42420 aut: Austrian Freedom Party
42420 aut: Austrian Freedom Party
42420 aut: Austrian Freedom Party
42420 aut: Austrian Freedom Party
42520 aut: Austrian People’s Party
42520 aut: Austrian People’s Party
42520 aut: Austrian People’s Party
42520 aut: Austrian People’s Party
42710 aut: Alliance for the Future of Austria
42710 aut: Alliance for the Future of Austria
43110 che: Green Party of Switzerland
43120 che: Green Liberal Party
43220 che: Swiss Labour Party
43320 che: Social Democratic Party of Switzerland
43420 che: Radical Democratic Party
43520 che: Conservative Christian Social Party
43530 che: Protestant People’s Party of Switzerland
43540 che: Christian Social Party
43711 che: Federal Democratic Union
43810 che: Swiss People’s Party
43811 che: Conservative Democratic Party of Switzer.
53021 irl: United Left Alliance
53110 irl: Green Party
53230 irl: Socialist Party
53320 irl: Labour Party
53420 irl: Progressive Democrats
53520 irl: Familiy of the Irish
53620 irl: Soldiers of Destiny
53951 irl: Ourselves Alone
61320 usa: Democratic Party
61620 usa: Republican Party
62110 can: Green Party
62320 can: New Democratic Party
62420 can: Liberal Party of Canada
62623 can: Conservative Party of Canada
62901 can: Quebec Bloc
63110 aus: Australian Greens
63320 aus: Australian Labor Party
63410 aus: Palmer United Party
63620 aus: Liberal Party of Australia

```
63710 aus: Katter's Australian Party
63810 aus: National Country Party
64110 nzl: Green Party of Aotearoa New Zealand
64320 nzl: New Zealand Labour Party
64420 nzl: ACT New Zealand
64421 nzl: United Future New Zealand
64422 nzl: Progressive Party
64620 nzl: New Zealand National Party
64621 nzl: New Zealand First Party
64901 nzl: Maori Party
64902 nzl: Mana Party
```

Comment: the party codes are the same as in the MARPOR datasets.

Date

```
-------------------------------------------------------------------------------
date                                                          Election date
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 range:  [199803,201309]              units:  1
         unique values:  32                       missing .: 0/235353
                  mean:    200741
             std. dev:    390.713
           percentiles:        10%       25%       50%       75%       90%
                            200209    200609    200803    201104    201111

-------------------------------------------------------------------------------
pos_corpus                                 Position of Qsentence in document
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 range:  [1,6836]                     units:  1
         unique values:  6713                     missing .:  235/235353
                  mean:   1113.06
             std. dev:   1049.95
           percentiles:        10%       25%       50%       75%       90%
                               113       324       824      1595      2426
```

Comment: Gives the position of the quasi-sentence within each document in the Manifesto Corpus. It can be used to merge the original verbatim of the quasi-sentences to the dataset. It is missing for 234 quasi-sentences from the Finnish National Coalition in 2007 and one quasi-sentence from the German Greens in 2013. For the crowd coding we have worked with the beta version of the Manifesto Corpus. The respective quasi- sentences were in the beta version, but are not in the publicly available Manifesto Corpus, because they were classified as text in margin by the Manifesto Project. The R-Script provided on the website makes it possible to add the verbatim from these quasi-sentences nonetheless.

```
-------------------------------------------------------------------------------
gs_1r                                      Was gold sentence in 1st round
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 range:  [0,1]                        units:  1
         unique values:  2                        missing .:  190/235353
```

Comment: The variable takes the value 0 if it was not a gold sentence and 1 if it was a gold sentence. Variable is missing for 190 quasi-sentences that were not included in the first coding round, but were added later and coded by the authors.

```
-------------------------------------------------------------------------------
gs_answer_1r                                  Test question answer 1st round
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 range:  [0,1]                        units:  1
         unique values:  2                    missing .: 234605/235353
          value labels: 0 Not immi/inti related
                        1 Immi/inti related

Comment: The variable gives the value of how the respective gold-sentence
was coded by the authors, i.e. 0 if it was classified as not related to
immigration or integration and 1 if it was regarded as related to one of
these. It is missing if a quasi-sentence was not a gold sentence in the
first round.
-------------------------------------------------------------------------------
gs_2r                                         Was gold sentence in 2nd round
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 range:  [0,1]                        units:  1
         unique values:  2                    missing .:  226393/235353
            tabulation:  Freq.  Value
                          8682  0
                           278  1
                        2.3e+05 .

Comment: The variable takes the value 0 if it was not a gold sentence and 1
if it was a gold sentence. Variable is missing if the quasi-sentence was
classified as not related to immigration or integration in the first round
and the quasi-sentence was thus not part of the second coding round.

-------------------------------------------------------------------------------
gs_answer_2q                                  Gold sentence answer 2nd round
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 label:  topic
                 range:  [1,2]                        units:  1
         unique values:  2                    missing .:  235075/235353
            tabulation:  Freq.         Numeric  Label
                           134         1  immigration
                           144         2  integration

Comment: See numeric label above. Variable is missing if quasi-sentence was
not a gold sentence in the second round.

-------------------------------------------------------------------------------
gs_answer_3q                                  Gold sentence answer direction
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 label:  direc
                 range:  [-1,1]                       units:  1
         unique values:  2                    missing .:  235075/235353
            tabulation:  Freq.        Numeric  Label
                           114        -1  sceptical
                           164         1  supportive
                        2.4e+05        .

Comment: See numeric label above. Variable is missing if quasi-sentence was
not a gold sentence.
```

```
-------------------------------------------------------------------------------
num_codings_1r                                      Number of coders 1st round
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 range:  [3,203]                      units:  1
         unique values:  151                  missing .:  190/235353
                  mean:   3.21572
             std. dev:   4.92441
           percentiles:        10%       25%       50%       75%       90%
                                 3         3         3         3         3

Comment: Number of codings per quasi-sentence in the first round. Minimum
three per sentence. If sentence was also used as a test then more codings.

-------------------------------------------------------------------------------
selection                          Immigration and/or integration related
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 range:  [0,1]                        units:  1
         unique values:  2                    missing .:  0/235353
            tabulation:  Freq.  Value
                        2.3e+05 0
                          8965  1

Comment: Variable takes the value of 1 if immigration or immigrant
integration related, and zero otherwise.

-------------------------------------------------------------------------------
certainty_selection                  Inter-coder agreement selection
-------------------------------------------------------------------------------
                  type:  numeric (double)
                 range:  [33.333333,100]              units:  1.000e-08
         unique values:  303                  missing .:  190/235353
                  mean:   97.6382
             std. dev:   8.93151
           percentiles:        10%       25%       50%       75%       90%
                               100       100       100       100       100

Comment: Degree of inter-coder agreement on the selection question. For
further details, please refer to the article accompanying this dataset
(Lehmann/Zobel 2018, EJPR).

-------------------------------------------------------------------------------
num_codings_2r                                      Number of coders 2nd round
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 range:  [5,84]                       units:  1
         unique values:  65                   missing .:  226393/235353
                  mean:   6.05234
             std. dev:   6.75641
           percentiles:        10%       25%       50%       75%       90%
                                 5         5         5         5         5

Comment: Number of codings per quasi-sentence in the second round. Minimum
five per sentence. If sentence was also used as a test then more codings.

-------------------------------------------------------------------------------
topic                                  Either immigration or integration
-------------------------------------------------------------------------------
                  type:  numeric (long)
                 label:  topic
```

```
                range:  [1,2]                       units:  1
        unique values:  2                       missing .:  226388/235353
           tabulation:  Freq.         Numeric  Label
                         3899         1  immigration
                         5066         2  integration
                       2.3e+05        .
```

Comments: See numeric labels above. Missing for quasi-sentences which are classified as not-related to immigration or integration.

---

certainty_topic — Inter-coder agreement topic

```
                type:  numeric (double)
                range:  [51.428571,100]             units:  1.000e-07
        unique values:  138                     missing .:  226393/235353
                 mean:   86.7369
            std. dev:   15.5884
          percentiles:        10%       25%       50%       75%       90%
                               60        80       100       100       100
```

Comment: Degree of inter-coder agreement on the topic question. For further details, please refer to the article accompanying this dataset (Lehmann/Zobel 2018, EJPR).

---

direction — Sceptical, supportive, or neutral

```
                type:  numeric (long)
               label:  direc
                range:  [-1,1]                      units:  1
        unique values:  3                       missing .:  226388/235353
           tabulation:  Freq.   Numeric  Label
                         2252        -1  sceptical
                          973         0  neutral
                         5740         1  supportive
                       2.3e+05        .
```

Comment: Gives the direction of a quasi-sentences as either sceptical, neutral or supportive. Missing for quasi-sentence which are classified as not-related to immigration or integration.

---

certainty_direction — Inter-coder agreement direction

```
                type:  numeric (double)
                range:  [37.5,100]                  units:  1.000e-07
        unique values:  138                     missing .:  226393/235353
                 mean:   81.4674
            std. dev:   19.5078
          percentiles:        10%       25%       50%       75%       90%
                               60        60        80       100       100
```

Comment: Degree of inter-coder agreement on the direction question. For further details, please refer to the article accompanying this dataset (Lehmann/Zobel 2018, EJPR).

---

manually_coded — Not coded by the crowd

```
                type:  numeric (long)
```

```
                range:  [0,1]                       units:  1
        unique values:  2                        missing .:  0/235353
           tabulation:  Freq.  Value
                       2.4e+05 0
                          221  1
```

Comment: 1 Australian and 189 German quasi-sentences were not coded by the crowd (because they were missing in the beta version of the Manifesto Corpus that was used to create the data for the Crowd coding) but were coded manually by the authors. In addition 31 quasi-sentences from Ireland were manually recoded as zero, because those covered emigration not immigration.

---

rn — Running number

```
                type:  numeric (long)
                range:  [1,235353]                  units:  1
        unique values:  235353                   missing .:  0/235353
                mean:    117677
            std. dev:   67940.7
          percentiles:        10%       25%       50%       75%       90%
                           23536     58839    117677    176515    211818
```

Comment: A running number sorting all quasi-sentences within each document into the order in which they appeared in the document.