UNIVERSITÀ DEGLI STUDI DI MILANO
FACOLTÀ DI SCIENZE POLITICHE,
ECONOMICHE E SOCIALI

Political communication and populist rhetoric, an analysis of Italian politicians in the digital arena.

By

RICCARDO RUTA

in partial fulfillment of the requirement
for the degree of ...
in Political, Economic and Social Sciences

07/22

Abstract

(the spacing is set to 1.5)

no more than 250 words for the abstract

- a description of the research question/knowledge gap – what we know and what we don't know
- how your research has attempted to fill this gap
- a brief description of the methods
- brief results
- key conclusions that put the research into a larger context

Contents

0.0.1 RICORSO ALLA RETORICA POPULISTA.

0.0.2 At the level of political parties, which ones make most use of populist rhetoric?

0.0.3 At the level of individual politicians, which ones make most use of populist rhetoric?

- Rooduijn & Pauwels: Rooduijn, M., and T. Pauwels. 2011. "Measuring Populism: Comparing Two Methods of Content Analysis." West European Politics 34 (6): 1272–1283.
- Decadri & Boussalis: Decadri, S., & Boussalis, C. (2020). Populism, party membership, and language complexity in the Italian chamber of deputies. Journal of Elections, Public Opinion and Parties, 30(4), 484-503.
- Grundl: Gründl J. Populist ideas on social media: A dictionary-based measurement of populist communication. New Media & Society. December 2020.
- Decadri & Boussalis + Grundl: this is simply a more extended version of the D&B dictionary, which also contains some terms taken from Grundl.

1 1) First step, import the words and create the dictionary

```
# import dictionaries file
dict <-  read_excel("data/populism_dictionaries.xlsx")
variable.names(dict)
```

```
## [1] "Rooduijn_Pauwels_Italian"
## [2] "Grundl_Italian_adapted"
## [3] "Decadri_Boussalis"
## [4] "Decadri_Boussalis_Grundl_People"
## [5] "Decadri_Boussalis_Grundl_Common Will"
## [6] "Decadri_Boussalis_Grundl_Elite"
```

```
# create the dictionary
Rooduijn_Pauwels_Italian <-
  dictionary(list(populism =
                   (dict$Rooduijn_Pauwels_Italian
                    [!is.na(dict$Rooduijn_Pauwels_Italian)])))

Grundl_Italian_adapted <-
  dictionary(list(populism =
                   dict$Grundl_Italian_adapted
                 [!is.na(dict$Grundl_Italian_adapted)]))

Decadri_Boussalis <-
  dictionary(list(populism =
                   dict$Decadri_Boussalis
                 [!is.na(dict$Decadri_Boussalis)]))

Decadri_Boussalis_Grundl <-
  dictionary(list(people =
                   dict$Decadri_Boussalis_Grundl_People
                 [!is.na(dict$Decadri_Boussalis_Grundl_People)],
                 common_will =
                   dict$`Decadri_Boussalis_Grundl_Common Will`
```

```
                    [!is.na(dict$`Decadri_Boussalis_Grundl_Common Will`)],
                    elite =
                      dict$Decadri_Boussalis_Grundl_Elite
                    [!is.na(dict$Decadri_Boussalis_Grundl_Elite)]))
```

2 2) Import the DFM prepared in previous steps and apply dictionaries

3 I) Decadri_Boussalis_Grundl

3.0.1 Level of sparsity

daily: 12.08%

weekly: 0.55%

monthly: 0%

```
# Daily Dictionary analysis with Decadri_Boussalis_Grundl on the whole dataset
dfm_dict1  <- dfm_lookup(dfm_weight, dictionary = Decadri_Boussalis_Grundl)
# Group by date
dfm_by_date1 <- dfm_group(dfm_dict1, groups= date)
dfm_by_date1
```

```
## Document-feature matrix of: 839 documents, 3 features (12.08% sparse) and 3 docva
##              features
## docs                people common_will      elite
##   2020-01-01 0.04761905  0           0.06666667
##   2020-01-02 0.42636838  0.05882353 0.39453879
```

```
##   2020-01-03 0.57511141  0.25000000 0.53910534
##   2020-01-04 1.35115485  0          0.28968254
##   2020-01-05 1.94643434  0.08333333 0.23852814
##   2020-01-06 0.75485528  0          1.16776316
## [ reached max_ndoc ... 833 more documents ]
```

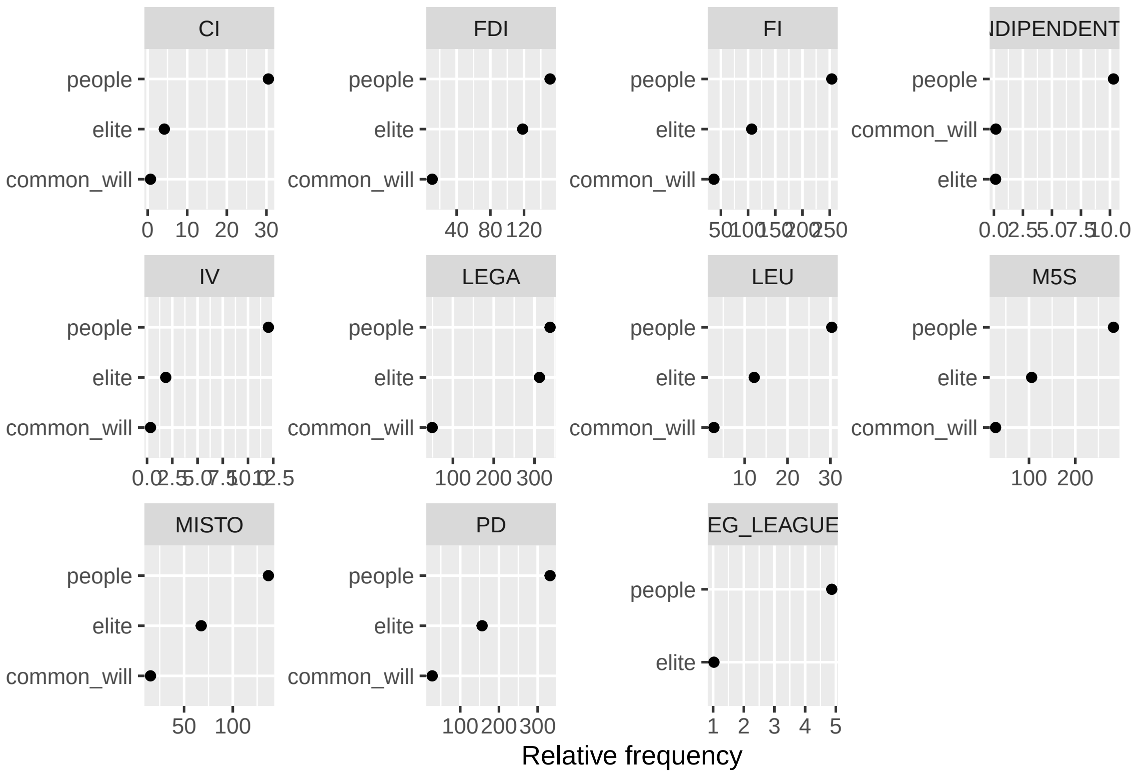
```r
# Group by week
dfm_by_week1 <- dfm_group(dfm_dict1, groups= week)
dfm_by_week1
```

```
## Document-feature matrix of: 121 documents, 3 features (0.55% sparse) and 1 docvar
##     features
## docs    people common_will     elite
##    1  4.346688   0.3921569  1.528521
##    2 10.917519   2.1764234  5.389965
##    3 16.217653   1.9467875  7.994374
##    4 17.996818   1.3673553  8.399551
##    5 17.661895   0.9670610  6.979555
##    6 11.257576   0.7009804 10.688212
## [ reached max_ndoc ... 115 more documents ]
```

```r
# Group by month
dfm_by_month1 <- dfm_group(dfm_dict1, groups= month)
dfm_by_month1
```

```
## Document-feature matrix of: 28 documents, 3 features (0.00% sparse) and 1 docvar
##     features
## docs   people common_will    elite
##    1 63.08421    6.720617 30.07071
```

```
##    2 51.95882    3.532448 36.73581
##    3 59.69107    3.065409 26.04866
##    4 51.97619    1.975715 38.92381
##    5 49.45054    1.012790 35.60162
##    6 43.57187    1.785750 39.61004
## [ reached max_ndoc ... 22 more documents ]
```



Looking at the populist rhetoric for each party divided into the 3 components people-centrism, anti-elitism and common-will, we note that the most frequent components is People-centrism.

4 II) Rooduijn_Pauwels_Italian

4.0.1 Level of sparsity

daily: 0.60%

weekly: 0.%

monthly: 0%

```
# Daily Dictionary analysis with Rooduijn_Pauwels_Italian on the whole dataset
dfm_dict2  <- dfm_lookup(dfm_weight, dictionary = Rooduijn_Pauwels_Italian)
# Group by date
dfm_by_date2 <- dfm_group(dfm_dict2, groups= date)
dfm_by_date2
```

```
## Document-feature matrix of: 839 documents, 1 feature (0.60% sparse) and 3 docvars
##             features
## docs            populism
##   2020-01-01 0.06666667
##   2020-01-02 0.34691975
##   2020-01-03 0.53910534
##   2020-01-04 0.28968254
##   2020-01-05 0.23852814
##   2020-01-06 1.16776316
## [ reached max_ndoc ... 833 more documents ]
```

```
# Group by week
dfm_by_week2 <- dfm_group(dfm_dict2, groups= week)
dfm_by_week2
```

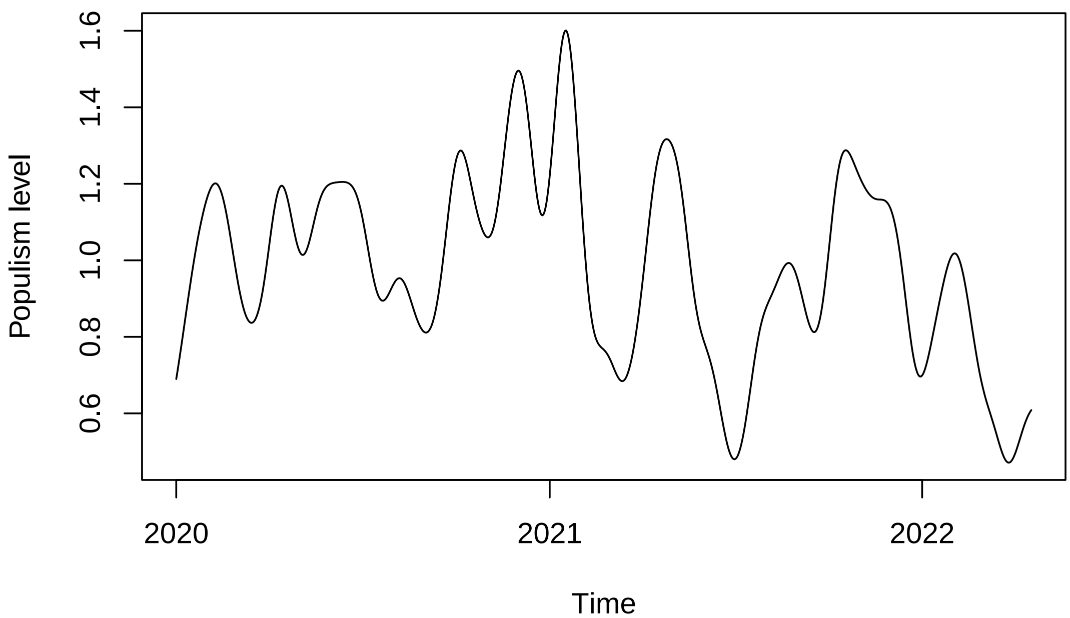
```
## Document-feature matrix of: 121 documents, 1 feature (0.00% sparse) and 1 docvar
```

```
##     features
## docs  populism
##    1  1.480902
##    2  5.140665
##    3  7.493528
##    4  7.872910
##    5  6.339158
##    6 10.353899
## [ reached max_ndoc ... 115 more documents ]
```

```r
# Group by month
dfm_by_month2 <- dfm_group(dfm_dict2, groups= month)
dfm_by_month2
```

```
## Document-feature matrix of: 28 documents, 1 feature (0.00% sparse) and 1 docvar.
##     features
## docs populism
##    1 28.10591
##    2 34.76596
##    3 24.91863
##    4 37.43421
##    5 32.79228
##    6 37.74417
## [ reached max_ndoc ... 22 more documents ]
```

4.1 General level of populism in time



4.2 Most populist party

```r
# Most populist party
dfm_dict2_tstat_party <- textstat_frequency(dfm_dict2, groups = party_id)
kable(dfm_dict2_tstat_party %>% slice_max(frequency, n = 20))
```

	feature	frequency	rank	docfreq	group
6	populism	303.9474786	1	1919	LEGA
10	populism	149.7512641	1	1671	PD
2	populism	113.7388243	1	1124	FDI
3	populism	98.6906136	1	941	FI
8	populism	87.6625041	1	1119	M5S
9	populism	60.9720255	1	669	MISTO
7	populism	11.7023384	1	175	LEU
1	populism	3.7116701	1	45	CI
5	populism	1.8540424	1	26	IV
11	populism	1.0264294	1	11	REG_LEAGUES
4	populism	0.0833333	1	1	INDIPENDENTE

4.3 Most populist politician

```
dict2_tstat_nome <- textstat_frequency(dfm_dict2, groups = nome)


kable(dict2_tstat_nome %>% slice_max(frequency, n = 20))
```

	feature	frequency	rank	docfreq	group
194	populism	42.115152	1	146	FERRERO Roberta
472	populism	15.910436	1	160	SGARBI Vittorio
341	populism	14.112659	1	77	MORANI Alessia
24	populism	13.999694	1	52	BALDELLI Simone
179	populism	13.821584	1	48	FAGGI Antonella
271	populism	13.095709	1	149	LANNUTTI Elio
217	populism	12.884799	1	39	FREGOLENT Sonia
450	populism	12.806346	1	64	RUSPANDINI Massimo
326	populism	12.518396	1	192	MELONI Giorgia
427	populism	12.257891	1	40	RIVOLTA Erica
106	populism	10.788399	1	68	CECCHETTI Fabrizio
283	populism	10.783981	1	108	LOLLOBRIGIDA Francesco
260	populism	10.778644	1	76	IEZZI Igor Giancarlo
230	populism	10.648954	1	155	GARNERO SANTANCHE' Daniela
303	populism	10.133849	1	78	MALAN Lucio
447	populism	9.885108	1	29	RUFA Gianfranco
455	populism	9.561830	1	93	SALVINI Matteo
360	populism	9.110910	1	105	NOBILI Luciano
35	populism	8.689617	1	57	BAZZARO Alex
501	populism	8.495460	1	32	TONELLI Gianni

5 III) Grundl_Italian_adapted

5.0.1 Level of sparsity

daily: 0.24%

weekly: 0.0%

monthly: 0%

```
# Daily Dictionary analysis with Grundl_Italian_adapted on the whole dataset
dfm_dict3  <- dfm_lookup(dfm_weight, dictionary = Grundl_Italian_adapted)
# Group by date
dfm_by_date3<- dfm_group(dfm_dict3, groups= date)
dfm_by_date3
```

```
## Document-feature matrix of: 839 documents, 1 feature (0.24% sparse) and 3 docvars
##             features
## docs            populism
##   2020-01-01 0
##   2020-01-02 0.16894258
##   2020-01-03 0.34545455
##   2020-01-04 0.08333333
##   2020-01-05 1.49644922
##   2020-01-06 1.64472860
## [ reached max_ndoc ... 833 more documents ]
```

```
# Group by week
dfm_by_week3 <- dfm_group(dfm_dict3, groups= week)
dfm_by_week3
```

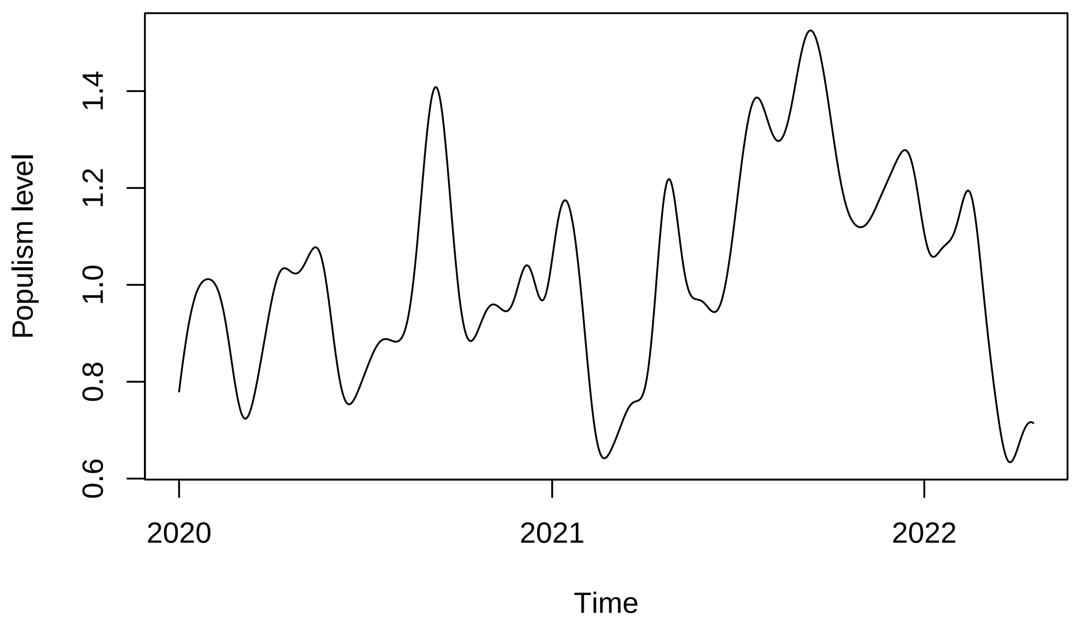
```
## Document-feature matrix of: 121 documents, 1 feature (0.00% sparse) and 1 docvar
##     features
## docs populism
##    1 2.094180
##    2 7.732182
##    3 7.349727
```

```
##    4 6.090898
##    5 7.887386
##    6 7.021892
## [ reached max_ndoc ... 115 more documents ]
```

```r
# Group by month
dfm_by_month3 <- dfm_group(dfm_dict3, groups= month)
dfm_by_month3
```

```
## Document-feature matrix of: 28 documents, 1 feature (0.00% sparse) and 1 docvar.
##     features
## docs populism
##    1 30.09665
##    2 26.23980
##    3 22.99661
##    4 32.36833
##    5 33.50214
##    6 21.44168
## [ reached max_ndoc ... 22 more documents ]
```

5.1 General level of populism in time



5.2 Most populist party

```
# Most populist party
dict_3_tstat_party <- textstat_frequency(dfm_dict3, groups = party_id)
kable(dict_3_tstat_party %>% slice_max(frequency, n = 20))
```

	feature	frequency	rank	docfreq	group
6	populism	225.678708	1	2075	LEGA
10	populism	153.269683	1	2017	PD
8	populism	133.053746	1	1724	M5S
3	populism	131.838292	1	1524	FI
2	populism	99.425177	1	1087	FDI
9	populism	86.092041	1	997	MISTO
7	populism	15.213765	1	231	LEU
1	populism	10.602522	1	157	CI
5	populism	2.559005	1	40	IV
4	populism	1.983671	1	31	INDIPENDENTE
11	populism	1.505044	1	22	REG_LEAGUES

5.3 Most populist politician

```
dict_3_tstat_nome <- textstat_frequency(dfm_dict3, groups = nome)


kable(dict_3_tstat_nome %>% slice_max(frequency, n = 20))
```

	feature	frequency	rank	docfreq	group
287	populism	23.033031	1	240	LANNUTTI Elio
210	populism	19.501980	1	110	FERRERO Roberta
562	populism	19.042283	1	131	VITO Elio
275	populism	16.483870	1	120	IEZZI Igor Giancarlo
494	populism	15.974269	1	184	SGARBI Vittorio
341	populism	11.063928	1	159	MELONI Giorgia
15	populism	10.731212	1	120	ANZALDI Michele
298	populism	10.659433	1	98	LOLLOBRIGIDA Francesco
74	populism	10.645964	1	97	BORGHI Claudio
476	populism	9.238862	1	122	SALVINI Matteo
248	populism	9.004085	1	139	GARNERO SANTANCHE' Daniela
96	populism	8.438949	1	103	CANGINI Andrea
546	populism	8.339166	1	106	URSO Adolfo
224	populism	8.162373	1	101	FONTANA Lorenzo
472	populism	7.850014	1	68	RUSPANDINI Massimo
44	populism	7.832168	1	120	BERGESIO Giorgio Maria
165	populism	7.565932	1	92	DE MARTINI Guido
141	populism	7.036558	1	43	CROSETTO Guido
446	populism	7.000320	1	47	RIVOLTA Erica
359	populism	6.861311	1	73	MORELLI Alessandro

6 IV) Decadri_Boussalis

6.0.1 Level of sparsity

daily: 0%

weekly: 0.0%

monthly: 0%

```r
# Daily Dictionary analysis with Decadri_Boussalis on the whole dataset
dfm_dict4  <- dfm_lookup(dfm_weight, dictionary = Decadri_Boussalis)
# Group by date
dfm_by_date4<- dfm_group(dfm_dict4, groups= date)
dfm_by_date4
```

```
## Document-feature matrix of: 839 documents, 1 feature (0.00% sparse) and 3 docvars
##             features
## docs           populism
##   2020-01-01 0.1142857
##   2020-01-02 0.8566215
##   2020-01-03 1.2364390
##   2020-01-04 1.6408374
##   2020-01-05 2.2375941
##   2020-01-06 1.9814420
## [ reached max_ndoc ... 833 more documents ]
```

```r
# Group by week
dfm_by_week4 <- dfm_group(dfm_dict4, groups= week)
dfm_by_week4
```

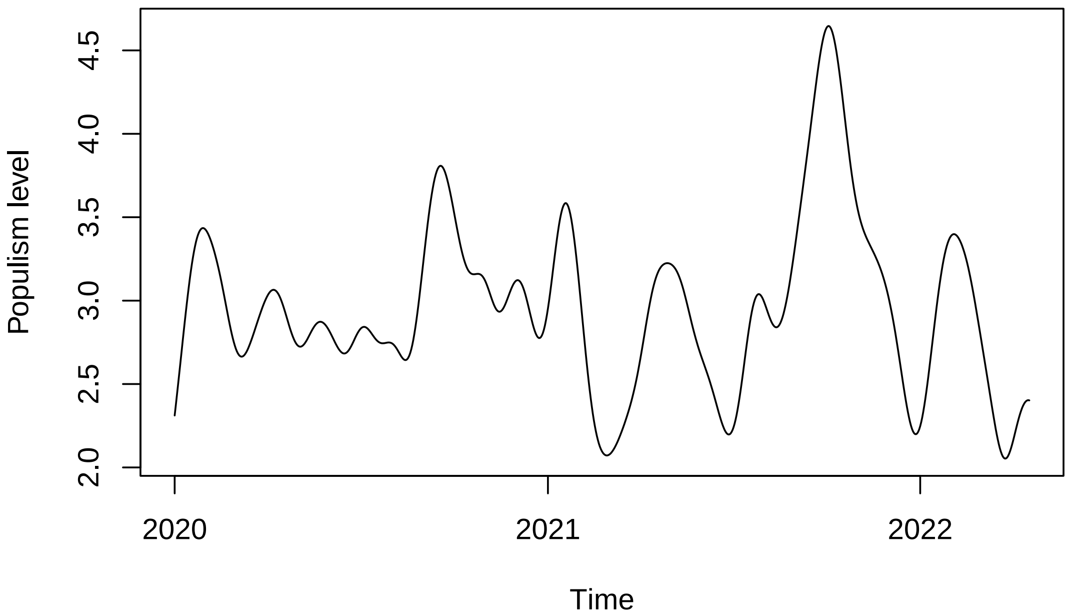
```
## Document-feature matrix of: 121 documents, 1 feature (0.00% sparse) and 1 docvar
##     features
## docs  populism
##    1  6.085778
##    2 16.762233
##    3 24.352100
```

```
##    4 26.321730
##    5 24.812520
##    6 22.284239
## [ reached max_ndoc ... 115 more documents ]
```

```
# Group by month
dfm_by_month4 <- dfm_group(dfm_dict4, groups= month)
dfm_by_month4
```

```
## Document-feature matrix of: 28 documents, 1 feature (0.00% sparse) and 1 docvar.
##     features
## docs populism
##    1 93.79618
##    2 88.79620
##    3 85.17899
##    4 90.99191
##    5 83.84470
##    6 82.69573
## [ reached max_ndoc ... 22 more documents ]
```

6.1 General level of populism in time



6.2 Most populist party

```
# Most populist party
dict_4_tstat_party <- textstat_frequency(dfm_dict4, groups = party_id)
kable(dict_4_tstat_party %>% slice_max(frequency, n = 20))
```

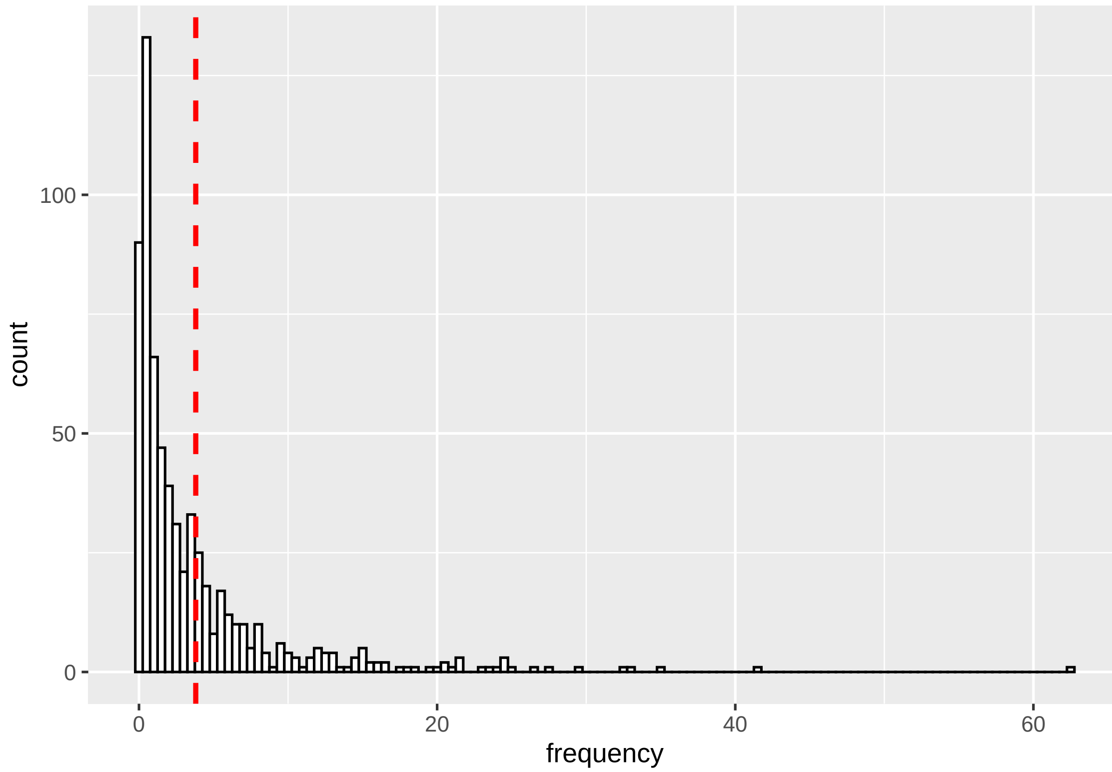
	feature	frequency	rank	docfreq	group
6	populism	651.348390	1	5672	LEGA
10	populism	493.532735	1	6417	PD
8	populism	376.966170	1	5178	M5S
3	populism	376.609606	1	4532	FI
2	populism	270.814483	1	2960	FDI
9	populism	202.466904	1	2463	MISTO
7	populism	44.919508	1	659	LEU
1	populism	35.105322	1	506	CI
5	populism	14.132863	1	197	IV
4	populism	10.615825	1	153	INDIPENDENTE
11	populism	6.122696	1	93	REG_LEAGUES

6.3 Most populist politician

```
dict_4_tstat_nome <- textstat_frequency(dfm_dict4, groups = nome)


kable(dict_4_tstat_nome %>% slice_max(frequency, n = 20))
```

	feature	frequency	rank	docfreq	group
236	populism	62.66405	1	282	FERRERO Roberta
560	populism	41.70723	1	443	SGARBI Vittorio
329	populism	34.85565	1	397	LANNUTTI Elio
391	populism	33.15912	1	496	MELONI Giorgia
344	populism	32.36912	1	358	LOLLOBRIGIDA Francesco
540	populism	29.61242	1	368	SALVINI Matteo
27	populism	27.44810	1	135	BALDELLI Simone
280	populism	26.74696	1	372	GARNERO SANTANCHE' Daniela
530	populism	24.85093	1	184	ROTONDI Gianfranco
68	populism	24.50676	1	252	BONACCINI Stefano
220	populism	24.35617	1	122	FAGGI Antonella
317	populism	24.31241	1	207	IEZZI Igor Giancarlo
128	populism	23.82148	1	195	CECCHETTI Fabrizio
585	populism	23.63509	1	327	TAJANI Antonio
80	populism	22.82617	1	240	BORGHI Claudio
161	populism	21.54784	1	158	CROSETTO Guido
39	populism	21.35229	1	202	BAZZARO Alex
47	populism	21.29380	1	318	BERGESIO Giorgio Maria
535	populism	20.92822	1	140	RUSPANDINI Massimo
365	populism	20.38171	1	185	MALAN Lucio

```
# TEST
ggplot(dict_4_tstat_nome, aes(x=frequency)) +
  geom_histogram(binwidth=.5, colour="black", fill="white") +
    geom_vline(aes(xintercept=mean(frequency, na.rm=T)),   # Ignore NA values for m
               color="red", linetype="dashed", size=1)
```


count
100
50
0
0
20
40
60
frequency

7 Inspect the Syuzhet dictionary

7.0.1 (http://saifmohammad.com/WebPages/lexicons.html)

```
head(get_sentiment_dictionary(dictionary = "nrc", language = "italian"),15)
```

```
##       lang          word sentiment value
## 1  italian          abba  positive     1
## 2  italian      capacità  positive     1
## 3  italian sopra citato  positive     1
## 4  italian      assoluto  positive     1
## 5  italian   assoluzione  positive     1
## 6  italian     assorbito  positive     1
## 7  italian    abbondanza  positive     1
## 8  italian    abbondante  positive     1
## 9  italian    accademico  positive     1
## 10 italian     accademia  positive     1
## 11 italian   accettabile  positive     1
## 12 italian accettazione  positive     1
## 13 italian   accessibile  positive     1
## 14 italian       encomio  positive     1
## 15 italian      alloggio  positive     1
```

7.0.2 Define function to make the text extracted from dataframe suitable for analysis

```
# Define function to make the text suitable for analysis
clean.text = function(x)
```

```r
{
  # tolower
  x = tolower(x)
  # remove rt
  x = gsub("rt", "", x)
  # remove at
  x = gsub("@\\w+", "", x)
  # remove punctuation
  x = gsub("[[:punct:]]", "", x)
  # remove numbers
  x = gsub("[[:digit:]]", "", x)
  # remove links http
  x = gsub("http\\w+", "", x)
  # remove tabs
  x = gsub("[ |\t]{2,}", "", x)
  # remove blank spaces at the beginning
  x = gsub("^ ", "", x)
  # remove blank spaces at the end
  x = gsub(" $", "", x)
  return(x)
}
```

7.1 First create the filtered dataframes

```r
# Create filtered dataframes
MELONI <- dataset %>% filter(nome  %like% "MELONI")
CONTE <- dataset %>% filter(nome  %like% "CONTE")
RENZI <- dataset %>% filter(nome  %like% "RENZI")
```

```
SALVINI <- dataset %>% filter(nome  %like% "SALVINI")
LETTA <- dataset %>% filter(nome  %like% "LETTA")
BERLUSCONI <- dataset %>% filter(nome  %like% "BERLUSCONI")
SPERANZA <- dataset %>% filter(nome  %like% "SPERANZA")
```

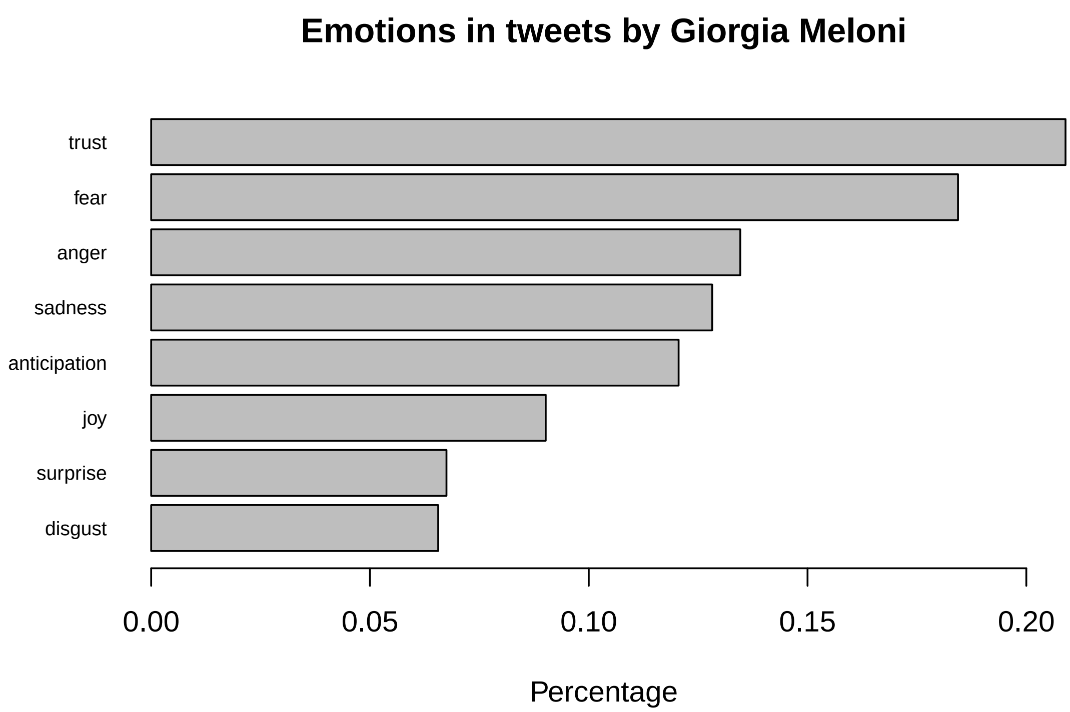
7.2 Then create nrc objects

```
# Create the nrc object
nrc_meloni <- get_nrc_sentiment(MELONI$tweet_testo, language="italian")
save(nrc_meloni,file="data/nrc_meloni.Rda")

nrc_conte <- get_nrc_sentiment(CONTE$tweet_testo, language="italian")
save(nrc_conte,file="data/nrc_conte.Rda")

nrc_renzi <- get_nrc_sentiment(RENZI$tweet_testo, language="italian")
save(nrc_renzi,file="data/nrc_renzi.Rda")

nrc_salvini <- get_nrc_sentiment(SALVINI$tweet_testo, language="italian")
save(nrc_salvini,file="data/nrc_salvini.Rda")

# NO DATA FOR LETTA
nrc_letta <- get_nrc_sentiment(LETTA$tweet_testo, language="italian")
save(nrc_letta,file="data/nrc_letta.Rda")

nrc_berlusconi <- get_nrc_sentiment(BERLUSCONI$tweet_testo, language="italian")
save(nrc_berlusconi, file="data/nrc_berlusconi.Rda")

nrc_speranza <- get_nrc_sentiment(SPERANZA$tweet_testo, language="italian")
```

```
save(nrc_speranza,file="data/nrc_speranza.Rda")
```

8 1) Giorgia Meloni - TRUST - GOVERNO

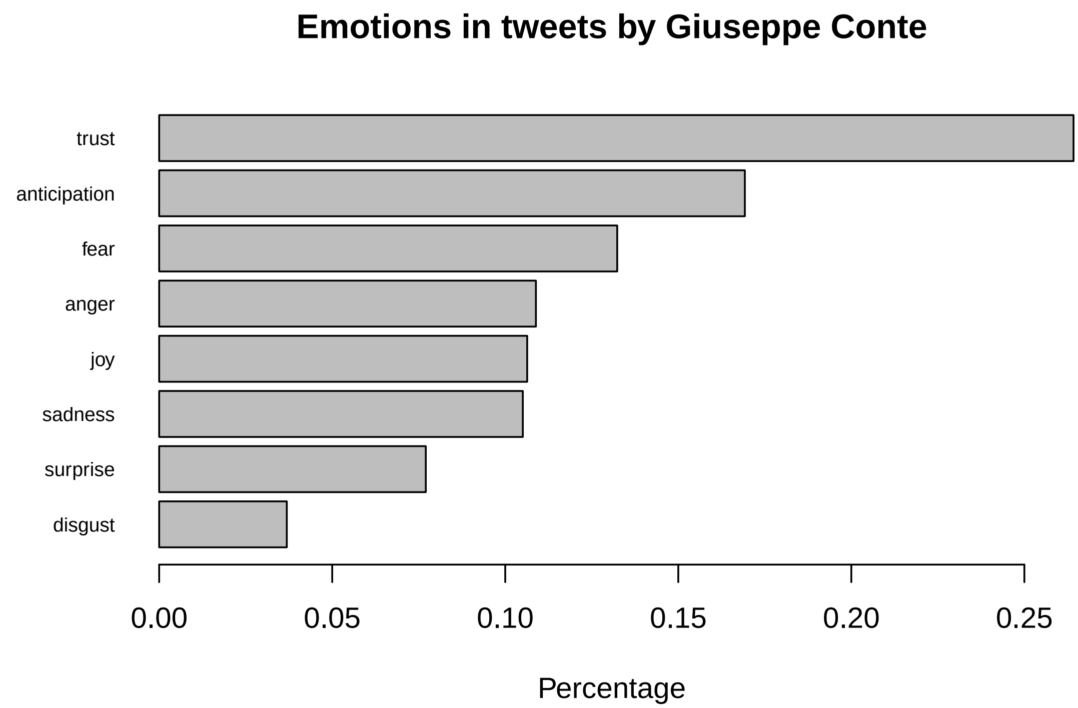
8.0.1 Plot the percentage of the emotion



8.0.2 Plot the wordcloud of emotions

Emotion Comparison Word Cloud for tweets by Giorgia Meloni

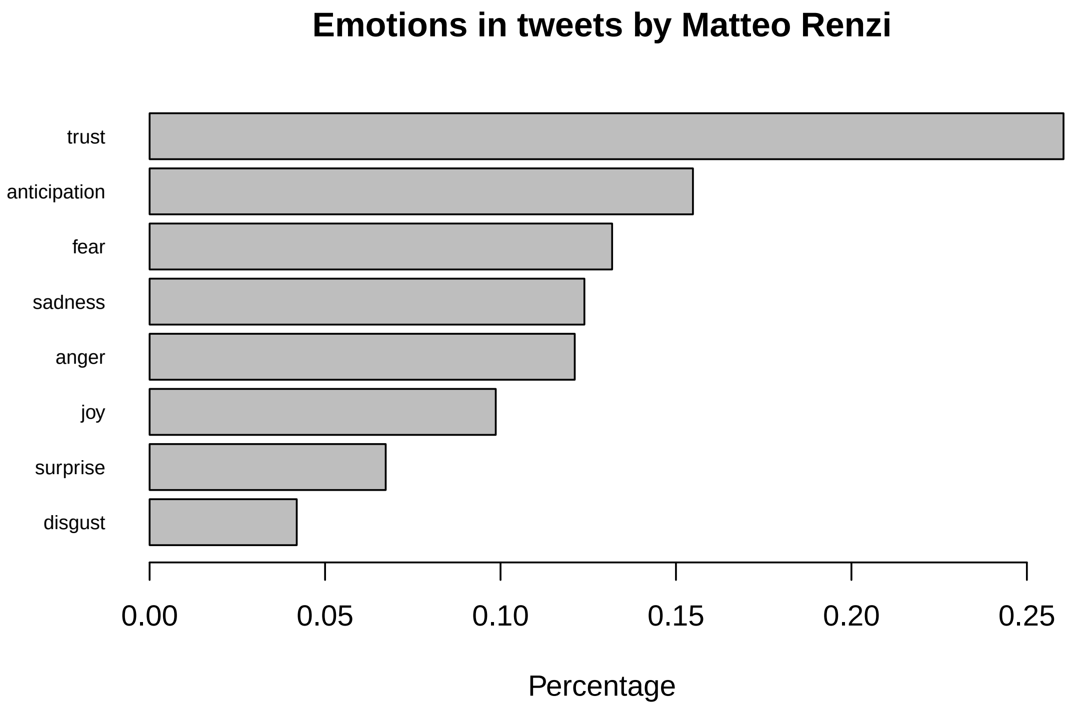
9 2) Conte - TRUST - LAVORO



Emotion Comparison Word Cloud for tweets by Giuseppe Conte

10 3) Renzi - TRUST - LAVORO

10.0.1 Plot the percentage of the emotion

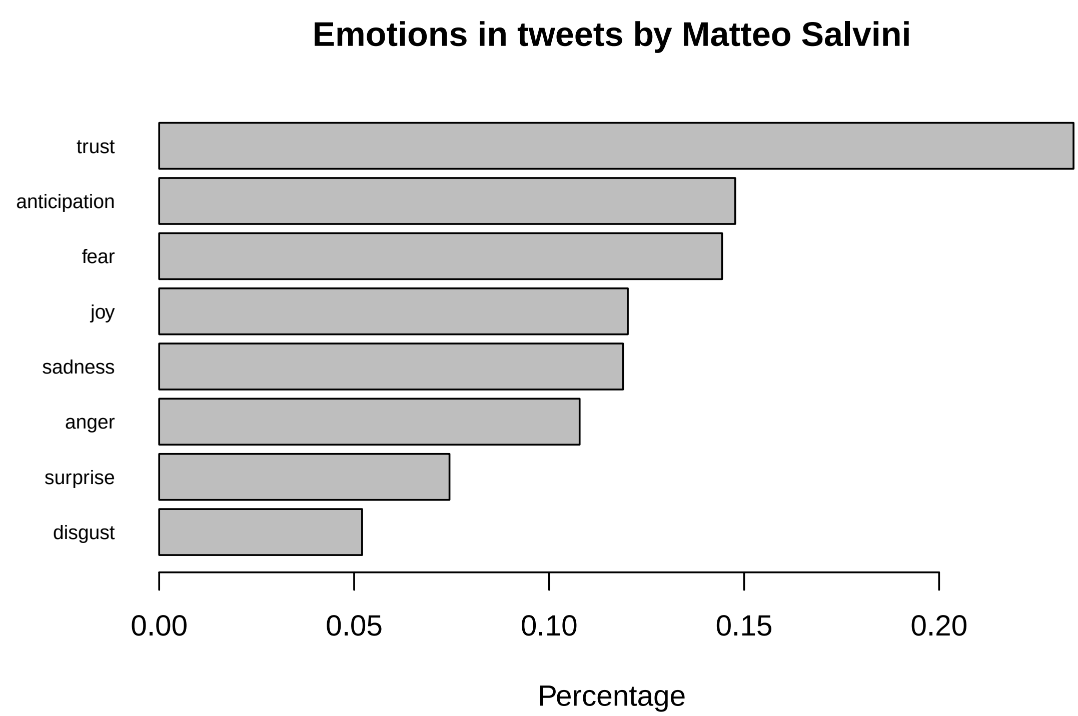


10.0.2 Plot the wordcloud of emotions

Emotion Comparison Word Cloud for tweets by Matteo Renzi

11 4) Salvini - TRUST - LAVORO

11.0.1 Plot the percentage of the emotion



11.0.2 Plot the wordcloud of emotions

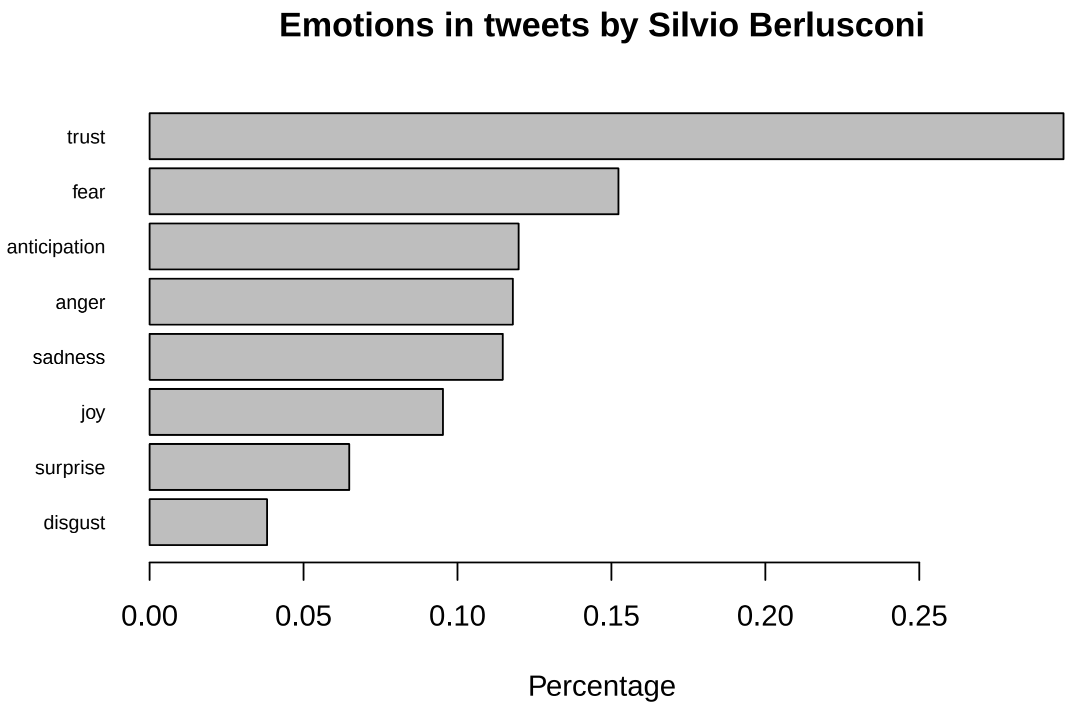
Emotion Comparison Word Cloud for tweets by Matteo Salvini

12 5) Letta

NO DATA FOR LETTA

13 6) Berlusconi - TRUST - GOVERNO

13.0.1 Plot the percentage of the emotion

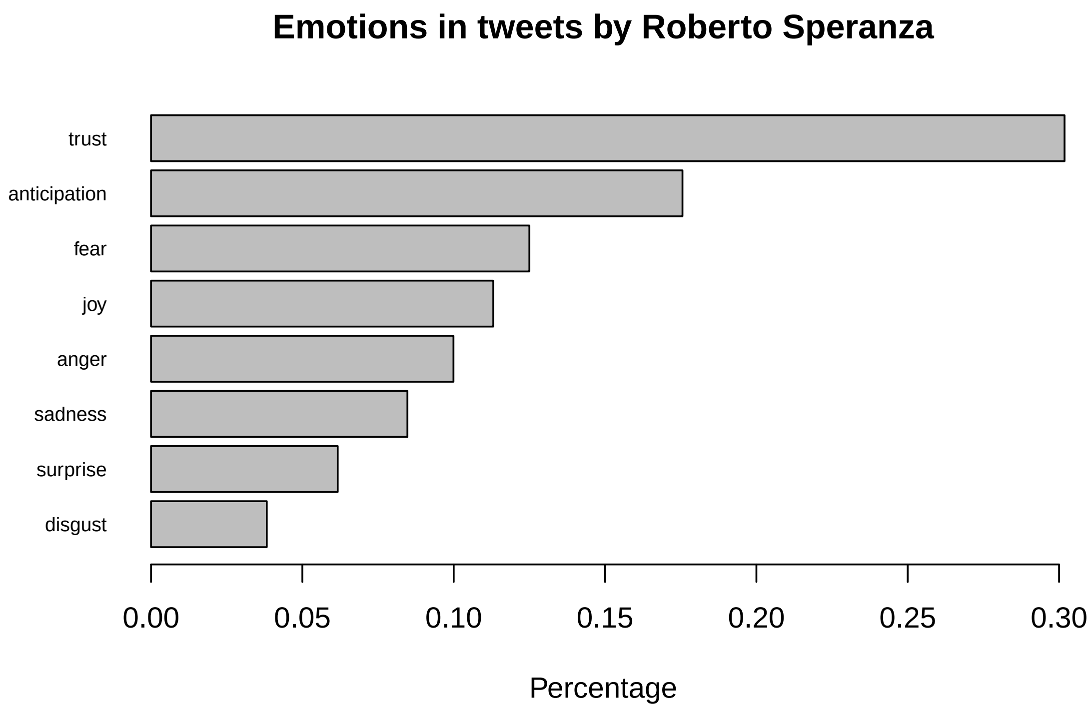


13.0.2 Plot the wordcloud of emotions

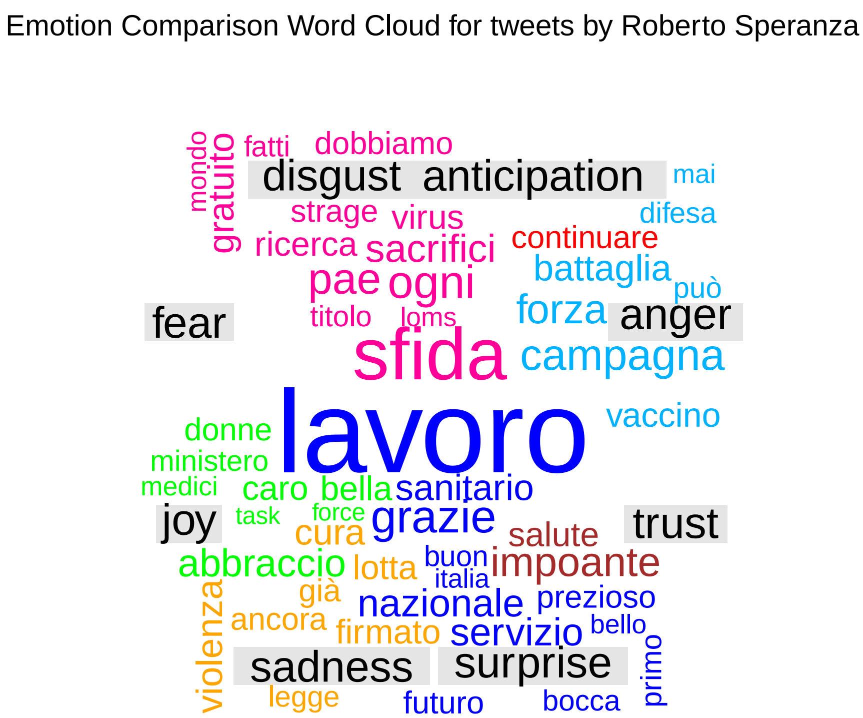
Emotion Comparison Word Cloud for tweets by Silvio Berlusconi

14 7) Speranza - TRUST - LAVORO

14.0.1 Plot the percentage of the emotion



14.0.2 Plot the wordcloud of emotions



14.1 PART I - CREATE THE DTM

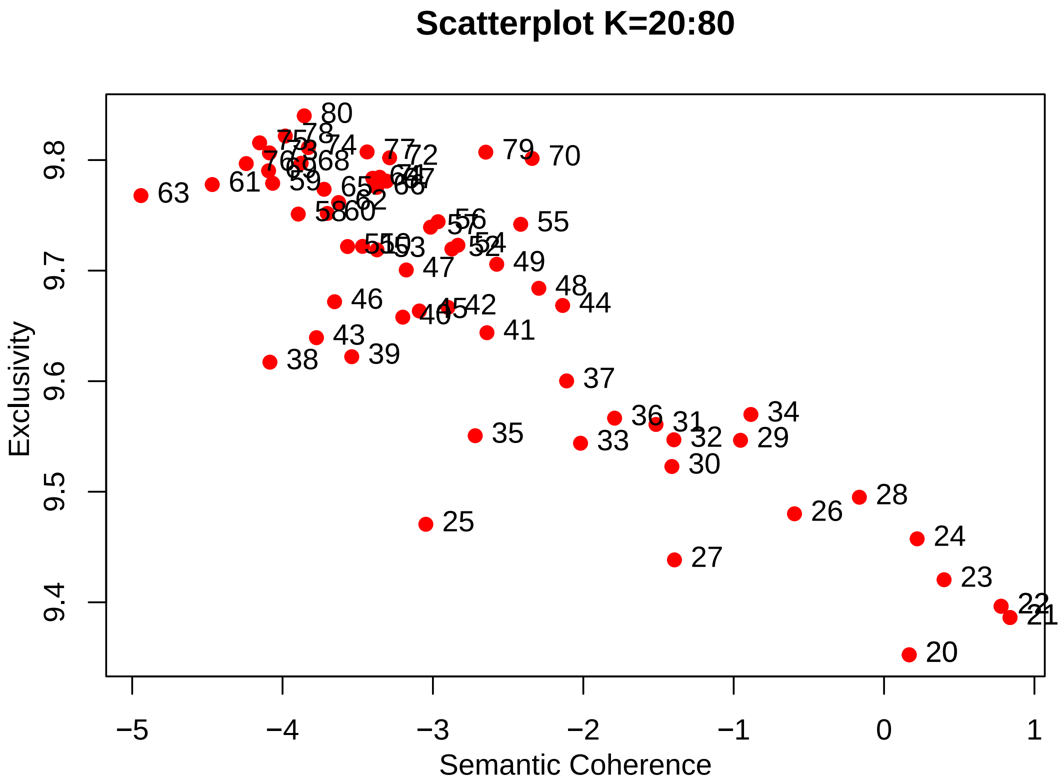
14.1.1 1) Convert the Document Feature Matrix (Dfm) in a Topic Model (Dtm)

14.2 PART II - FIND THE BEST NUMBER OF TOPICS K

14.2.1 1) First try K = 20:80

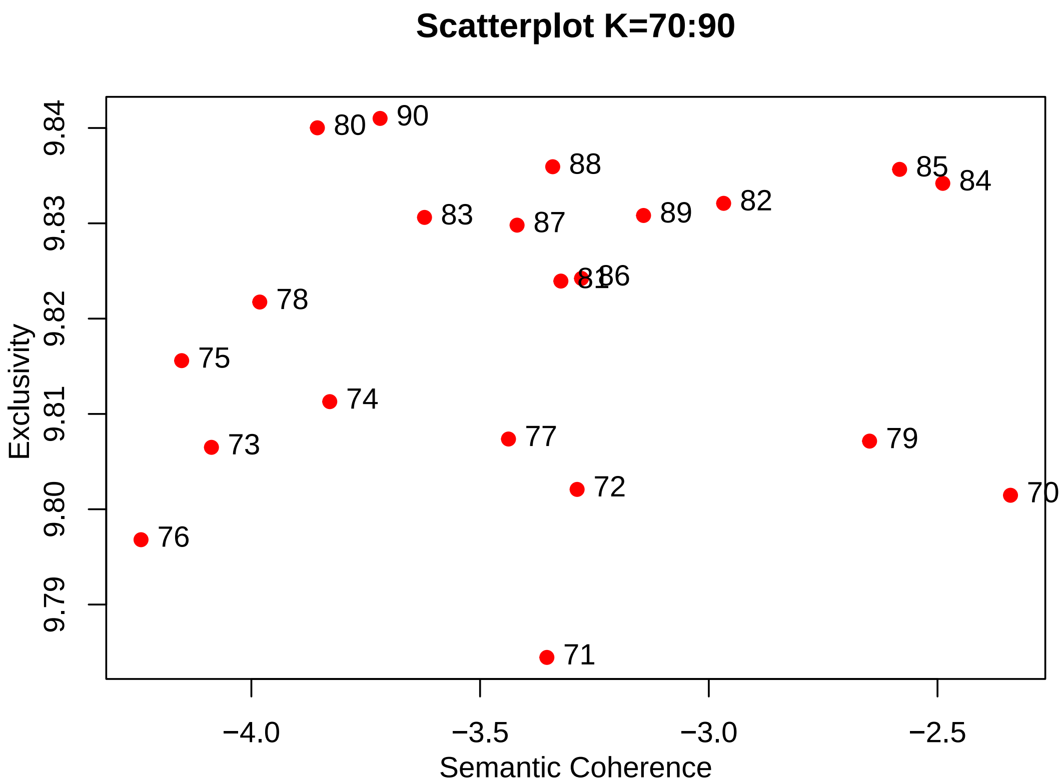
topic	coherence	exclusivity
20	0.1674464	9.352493
21	0.8376643	9.386225
22	0.7783789	9.396400
23	0.3991629	9.420433
24	0.2202299	9.457425
25	-3.0470840	9.470567

	topic	coherence	exclusivity
59	78	-3.9815896	9.821740
60	79	-2.6484726	9.807149
61	80	-3.8557027	9.840021
62	20	0.1674464	9.352493
63	21	0.8376643	9.386225
64	22	0.7783789	9.396400



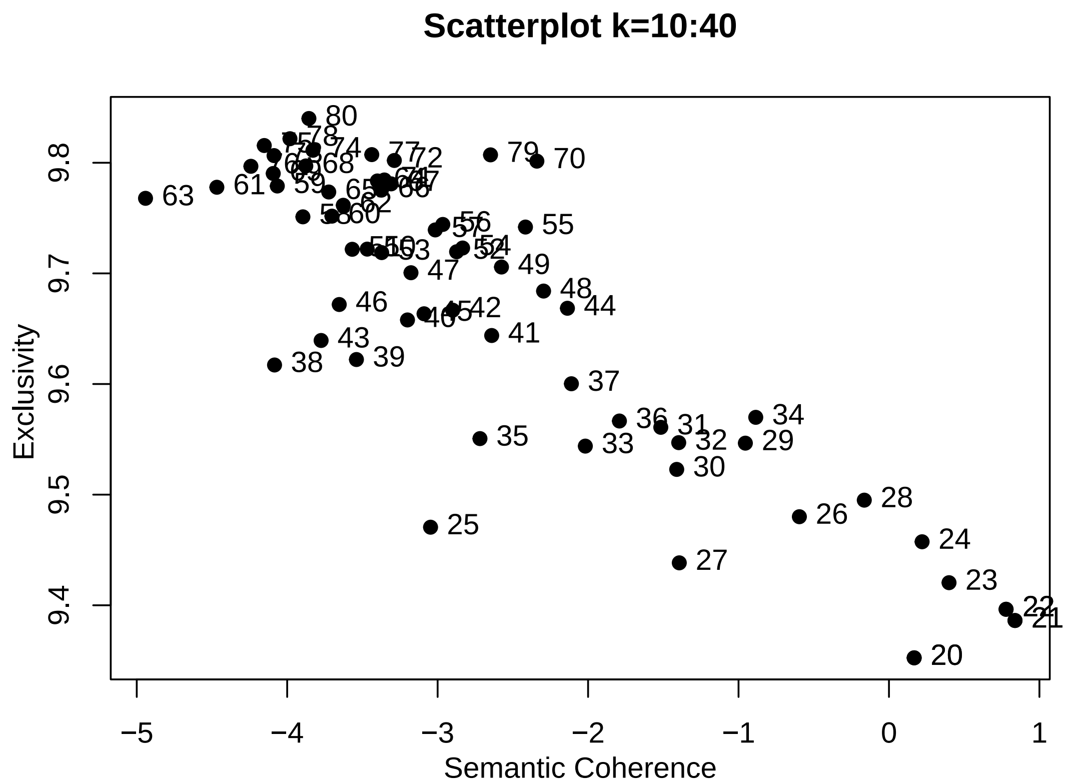
From this first try the best k seems to be 34

14.2.2 2) Second try K = 70:90



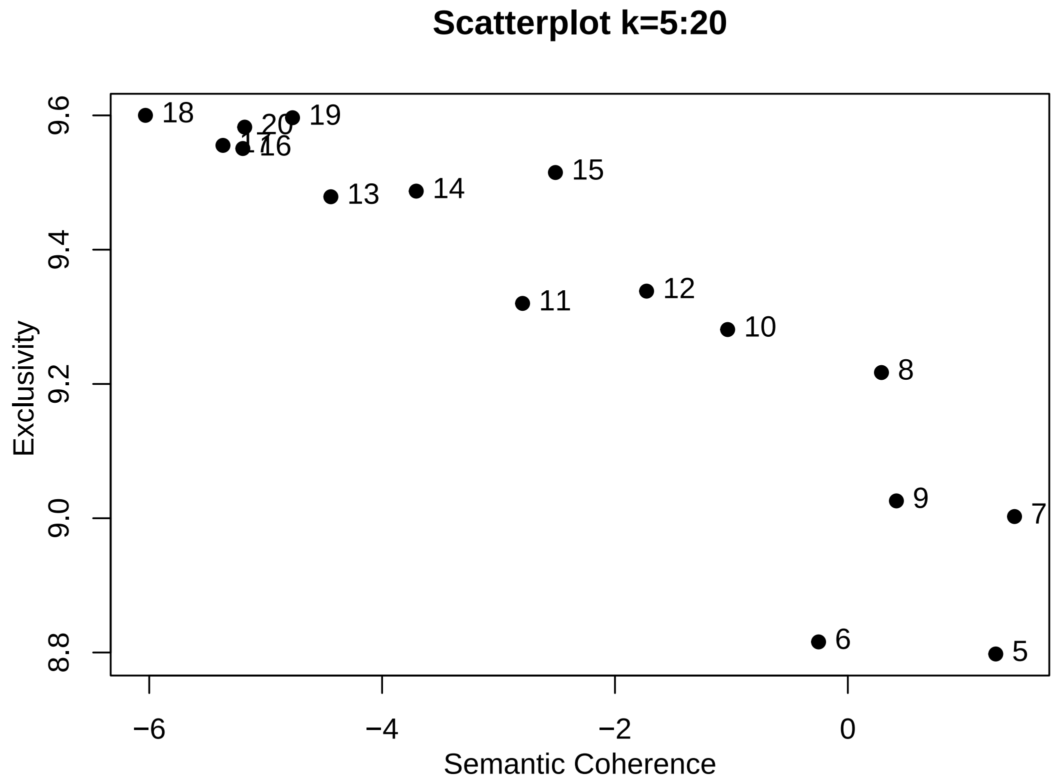
In this case 84 seems better, but the plot is very dispersive

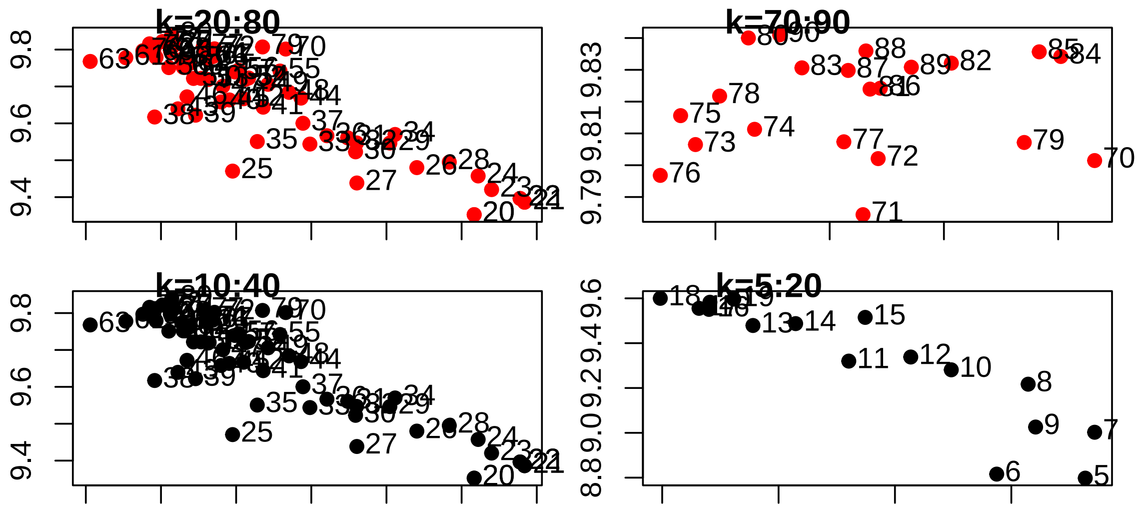
14.2.3 3) Third try K = 10:40 with iteration = 1000

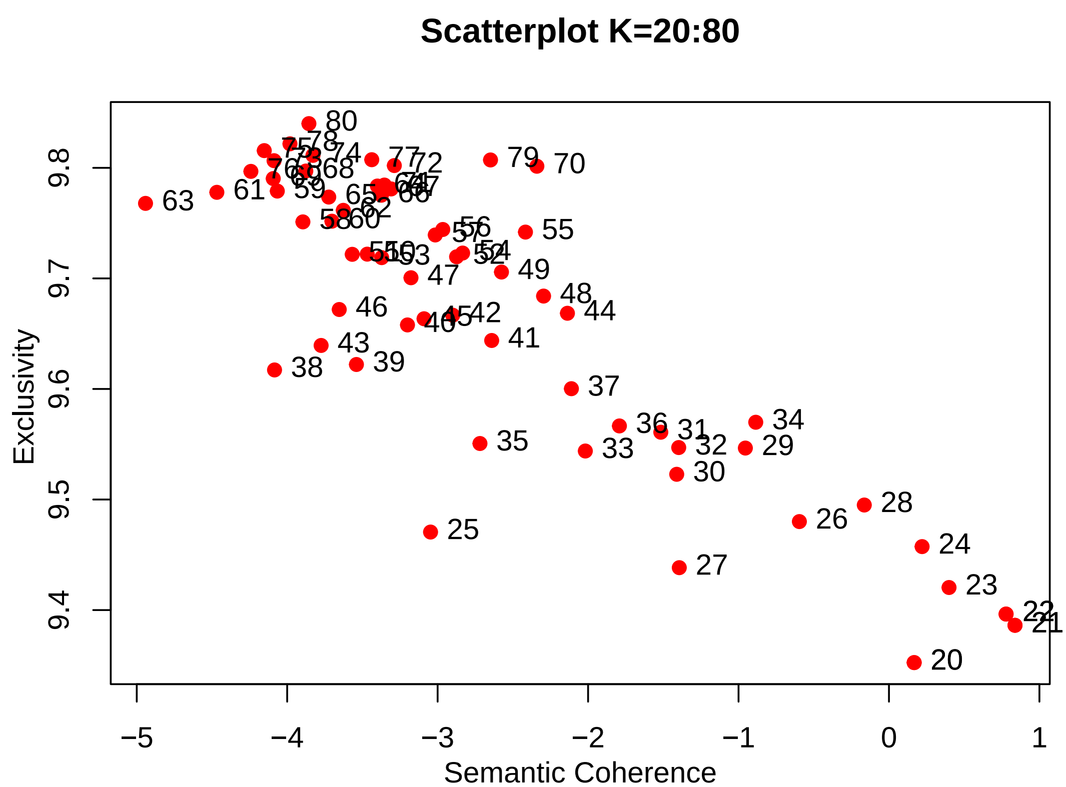


In this iteration the best choice can be 28 but the general level of coherence has fallen very low

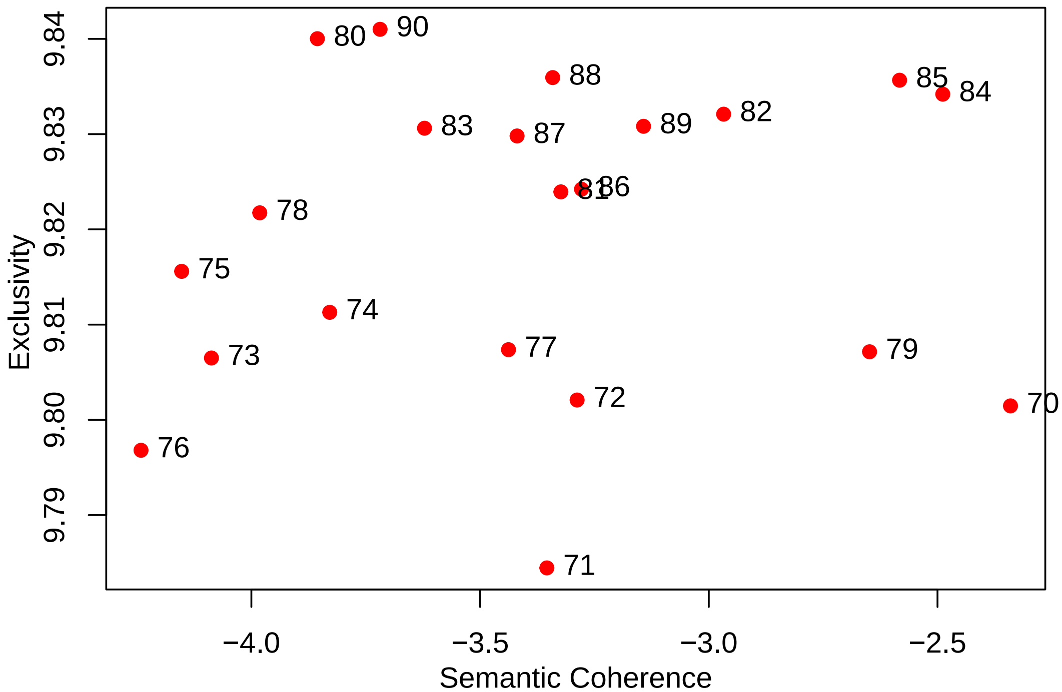
14.2.4 4) Fourth try k = 5:20 iteration n = 2000



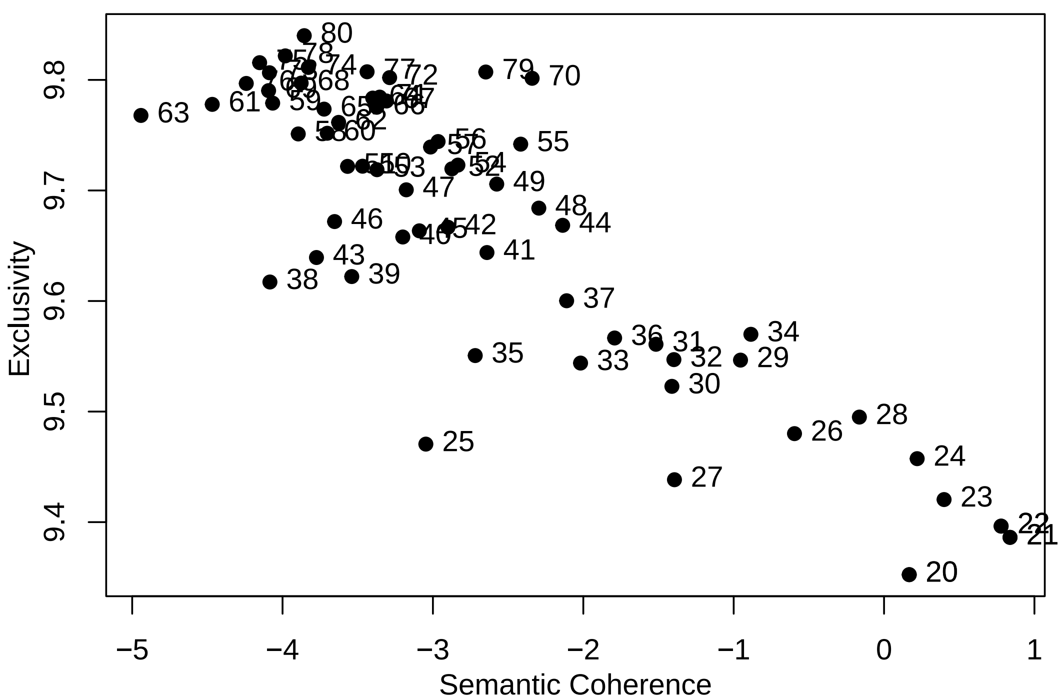

k=20:80
k=70:90
k=10:40
k=5:20

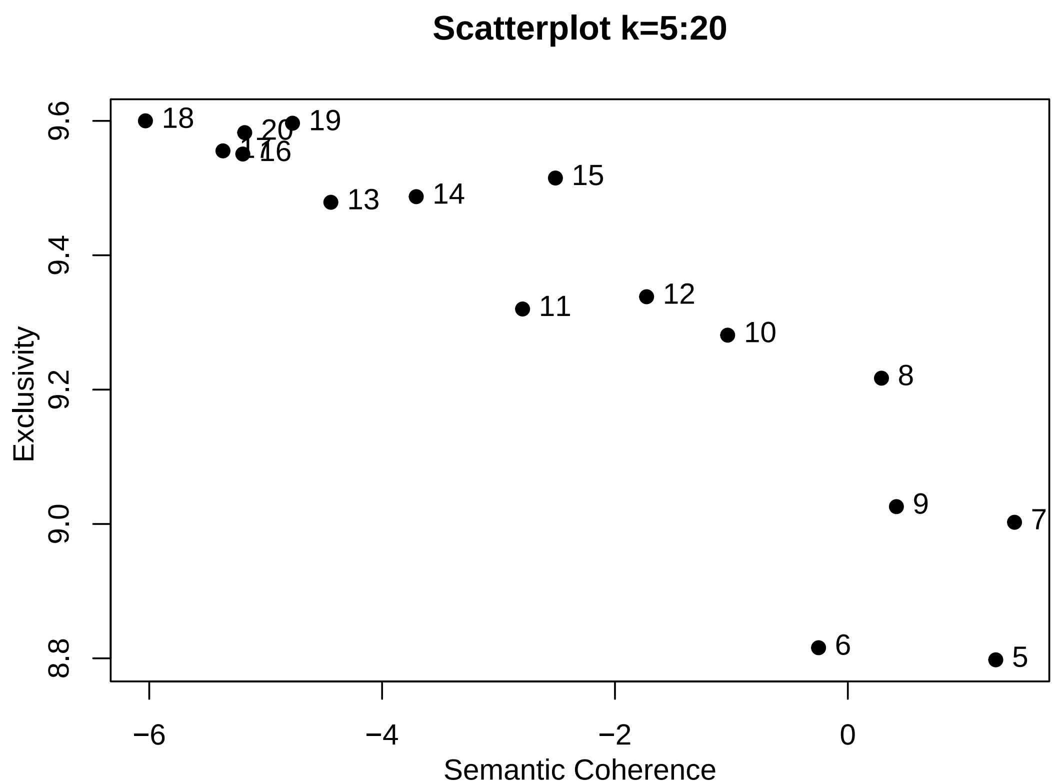

Scatterplot K=20:80
Exclusivity
Semantic Coherence

Scatterplot K=70:90



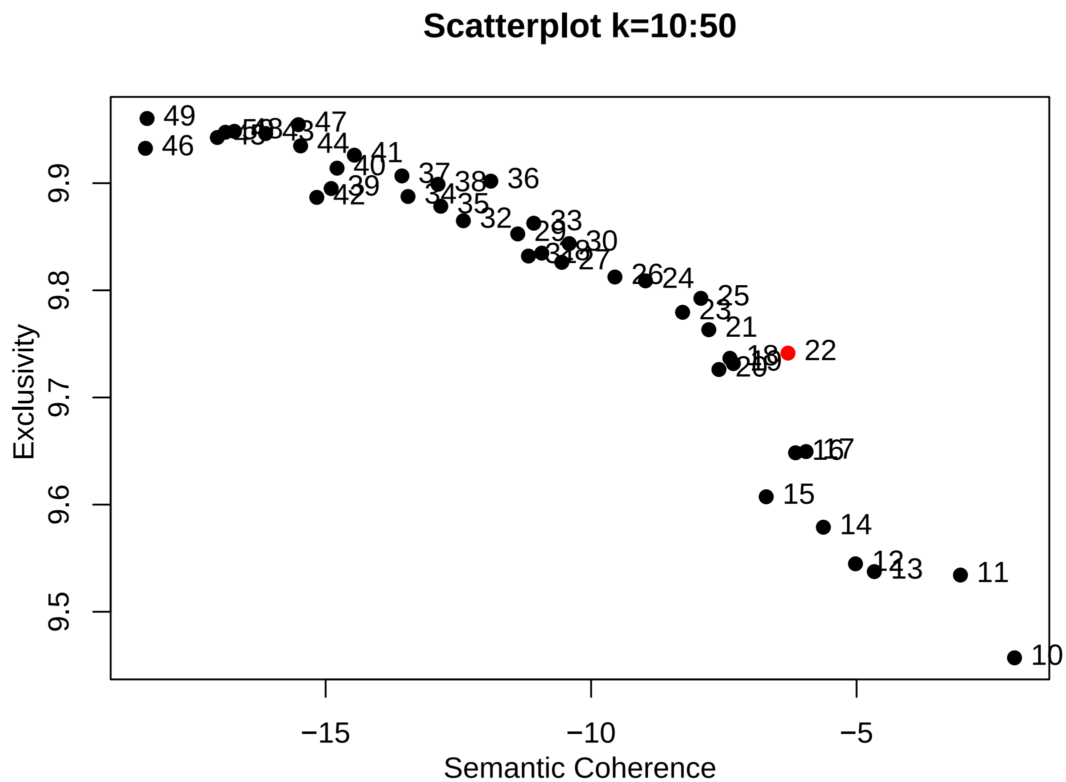
Scatterplot k=10:40





After all these tests, I believe that 10 is the best choice because it achieves good levels of coherence and exclusivity and is consistent with the choice of looking for substantive policy themes

14.2.5 5) Fifth try k = 10:50 iteration n = 1000



14.3 PART III - ANALISYS OF THE TOPICS

14.3.1 First try with k = 30

14.3.2 Here i extract the most important terms from the model

Top terms 01	Top terms 02	Top terms 03	Top terms 04	Top terms 05	Top terms 06	Top terms 07	Top terms 08	Top terms 09	Top terms 10
#afghanistan	president	forza	governo		#iostoconsalvini	#draghi	maggio	the	guerra
#tokyo2020	#quirinal	grand	italiani	donn	governo	draghi	governo	@fratelliditalia	#ucraina
talebani	repubblica	buon	+	via	luglio	governo	#decretorilancio	of	pace
	#presidentedellarepubblica	pd	lavoro	violenza	#recoveryfund	#governodraghi	#fase2	to	ucraina
agosto	#quirinale2022	l'italia	#covid19	giornata	president	lavoro	lavoro	and	putin
pass	draghi	politica	crisi		#cont	paes	ministro	violenza	donn
	gennaio		impres	minacc	#salvini	president	#bonafed	donn	marzo
@stampasgarbi	quirinal	c'è	diretta		legg	@fratelliditalia	#silviaromano	novembr	
grazi	grand	casa	momento	mondo	cont	buon	#recoveryfund	covid	russia
afghanistan	#mattarella	@pdnetwork	l'italia	pensiero	l'italia	l'italia	ripartir	#morradimett	ucraino

Top terms 11	Top terms 12	Top terms 13	Top terms 14	Top terms 15	Top terms 16	Top terms 17	Top terms 18	Top terms 19	Top terms 20
#dpcm	pass	maggio	pass	giugno	giugno	governo	governo	#coronavirus	lavoro
governo	sindaco	april	draghi	#2giugno	#primalitalia	ministro	cont	#mes	paes
#iostoconsalvini	green	vaccinal	natal	scuola	roma	paes	#crisidigoverno	#covid19	italia
ottobr	#greenpass	coprifuoco	green	minacc	bocca	c'è	crisi	#cont	donn
#cont	settembr	@stampasgarbi	vaccinati	+	luglio	cittadini	#cont	april	giornata
#mes	candidato	@fratelliditalia	dicembr	governo	piazza	president	paes	#forzalombardia	commission
covid	città	#nocoprifuoco		#cont	@stampasgarbi	bene	maggioranza	mes	l'italia
@fratelliditalia	draghi	#pnrr	@fratelliditalia	paes	forza		president	liquidità	italiani
de	piazza	pandemia	covid	@luigidimaio		grazi	#governo	ripartir	#lega
jole	roma	draghi	@fattoquotidiano	cont	draghi	parlamento	@fratelliditalia	#fase2	insiem

Top terms 21	Top terms 22	Top terms 23	Top terms 24	Top terms 25	Top terms 26	Top terms 27	Top terms 28	Top terms 29	Top terms 30
settembr	governo	#sanremo2022	#coronavirus	grazi	natal	vaccini	governo	pass	#referendumgiustizia
#iovotono	#iostoconsalvini	febbraio	grazi	anni	cont	donn	agosto	green	giustizia
elettoral	#oggivotolega	draghi	misur	lavoro	governo	buon	anni	ottobr	pass
#processateanchem	#salvini	green	l'emergenza	grand	bilancio	@pdnetwork	vittim	#ddlzan	luglio
#referendum	salvini	pass	momento	diritti	@fratelliditalia	marzo	settembr	roma	#greenpass
parlamentari	#borgonzonipresident	#ucraina	emergenza	via	dicembr	@fratelliditalia	#iovotono	sindaco	gazebo
voto	#prescrizion	parlamento	coronavirus	legg	#mes	auguri	@fratelliditalia	legg	#ddlzan
@fratelliditalia	@fratelliditaiia	@fratelliditalia	casa	president	#natal	draghi	covid		riforma
scuola	#emiliaromagna	guerra	decreto	giovani	italiani	vaccinal	bonus	+	firm
referendum	#m5s	#greenpass	#iorestoacasa	città	mes	vaccino	scuola	@forza_italia	draghi

14.3.3

COMMENTHERE

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
titles	1	2	3	4	5	6	7	8	9	10
	#afghanistan	president	forza	governo		#iostoconsalvini	#draghi	maggio	the	guerra
	#tokyo2020	#quirinal	grand	italiani	donn	governo	draghi	governo	@fratelliditalia	#ucraina
	talebani	repubblica	buon	+	via	luglio	governo	#decretorilancio	of	pace
		#presidentedellarepubblica	pd	lavoro	violenza	#recoveryfund	#governodraghi	#fase2	to	ucraina
	agosto	#quirinale2022	l'italia	#covid19	giornata	president	lavoro	lavoro	and	putin
	pass	draghi	politica	crisi		#cont	paes	ministro	violenza	donn
		gennaio		impres	minacc	#salvini	president	#bonafed	donn	marzo
	@stampasgarbi	quirinal	c'è	diretta		legg	@fratelliditalia	#silviaromano	novembr	
	grazi	grand	casa	momento	mondo	cont	buon	#recoveryfund	covid	russia
	afghanistan	#mattarella	@pdnetwork	l'italia	pensiero	l'italia	l'italia	ripartir	#morradimett	ucraino

	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15	Topic 16	Topic 17	Topic 18	Topic 19	Topic 20
titles	11	12	13	14	15	16	17	18	19	20
	#dpcm	pass	maggio	pass	giugno	giugno	governo	governo	#coronavirus	lavoro
	governo	sindaco	april	draghi	#2giugno	#primalitalia	ministro	cont	#mes	paes
	#iostoconsalvini	green	vaccinal	natal	scuola	roma	paes	#crisidigoverno	#covid19	italia
	ottobr	#greenpass	coprifuoco	green	minacc	bocca	c'è	crisi	#cont	donn
	#cont	settembr	@stampasgarbi	vaccinati	+	luglio	cittadini	#cont	april	giornata
	#mes	candidato	@fratelliditalia	dicembr	governo	piazza	president	paes	#forzalombardia	commission
	covid	città	#nocoprifuoco		#cont	@stampasgarbi	bene	maggioranza	mes	l'italia
	@fratelliditalia	draghi	#pnrr	@fratelliditalia	paes	forza		president	liquidità	italiani
	de	piazza	pandemia	covid	@luigidimaio		grazi	#governo	ripartir	#lega
	jole	roma	draghi	@fattoquotidiano	cont	draghi	parlamento	@fratelliditalia	#fase2	insiem

	Topic 21	Topic 22	Topic 23	Topic 24	Topic 25	Topic 26	Topic 27	Topic 28	Topic 29	Topic 30
titles	21	22	23	24	25	26	27	28	29	30
	settembr	governo	#sanremo2022	#coronavirus	grazi	natal	vaccini	governo	pass	#referendumgiustizia
	#iovotono	#iostoconsalvini	febbraio	grazi	anni	cont	donn	agosto	green	giustizia
	elettoral	#oggivotolega	draghi	misur	lavoro	governo	buon	anni	ottobr	pass
	#processateanchem	#salvini	green	l'emergenza	grand	bilancio	@pdnetwork	vittim	#ddlzan	luglio
	#referendum	salvini	pass	momento	diritti	@fratelliditalia	marzo	settembr	roma	#greenpass
	parlamentari	#borgonzonipresident	#ucraina	emergenza	via	dicembr	@fratelliditalia	#iovotono	sindaco	gazebo
	voto	#prescrizion	parlamento	coronavirus	legg	#mes	auguri	@fratelliditalia	legg	#ddlzan
	@fratelliditalia	@fratelliditaiia	@fratelliditalia	casa	president	#natal	draghi	covid		riforma
	scuola	#emiliaromagna	guerra	decreto	giovani	italiani	vaccinal	bonus	+	firm
	referendum	#m5s	#greenpass	#iorestoacasa	città	mes	vaccino	scuola	@forza_italia	draghi

14.3.4

COMMENTHERE

14.3.5 Repeat the search using a much lower K.

14.3.6 K = 10

Top terms 01	Top terms 02	Top terms 03	Top terms 04	Top terms 05	Top terms 06	Top terms 07	Top terms 08	Top terms 09	Top terms 10
guerra	grazie	#coronavirus	governo	draghi	governo	green	giustizia	#iostoconsalvini	governo
presidente	italia	#covid19	paese	#draghi	conte	via	grazie	salvini	italiani
#ucraina	legge	governo	presidente	lavoro	@fratelliditalia	draghi	#referendumgiustizia	governo	#covid19
draghi	donne	momento	lavoro	@fratelliditalia	the	@fattoquotidiano	#tokyo2020	elettorale	l'italia
ucraina	grande	misure	politica	vaccini	#conte	#greenpass	#greenpass	#salvini	@fratelliditalia
#quirinale	insieme	emergenza	italiani	pandemia	italiani	vaccinati	luglio	settembre	imprese
@fratelliditalia	via	grazie	l'italia	governo	#covid19	@fratelliditalia	@stampasgarbi	#iovotono	lavoro
putin	commissione	imprese	grande	vaccinale	#governo	sindaco	gazebo	#lega	#recoveryfund
marzo	libertà	decreto	grazie	#covid19	covid	ottobre	riforma	@matteosalvinimi	conte
russia	mondo	l'emergenza	parlamento	paese	pandemia	lavoro	green	@pdnetwork	#conte

14.3.7

COMMENTHERE

14.3.8 Titles:

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
titles	1	2	CORONAVIRUS	SOSTEGNI ECONOMICI	LEGA	DIRITTI	GOVERNO	INFORMAZIONE GIORNALISTICA	9	10
	guerra	grazie	#coronavirus	governo	draghi	governo	green	giustizia	#iostoconsalvini	governo
	presidente	italia	#covid19	paese	#draghi	conte	via	grazie	salvini	italiani
	#ucraina	legge	governo	presidente	lavoro	@fratelliditalia	draghi	#referendumgiustizia	governo	#covid19
	draghi	donne	momento	lavoro	@fratelliditalia	the	@fattoquotidiano	#tokyo2020	elettorale	l'italia
	ucraina	grande	misure	politica	vaccini	#conte	#greenpass	#greenpass	#salvini	@fratelliditalia
	#quirinale	insieme	emergenza	italiani	pandemia	italiani	vaccinati	luglio	settembre	imprese
	@fratelliditalia	via	grazie	l'italia	governo	#covid19	@fratelliditalia	@stampasgarbi	#iovotono	lavoro
	putin	commissione	imprese	grande	vaccinale	#governo	sindaco	gazebo	#lega	#recoveryfund
	marzo	libertà	decreto	grazie	#covid19	covid	ottobre	riforma	@matteosalvinimi	conte
	russia	mondo	l'emergenza	parlamento	paese	pandemia	lavoro	green	@pdnetwork	#conte

14.3.9

COMMENTHERE

14.3.10 Repeat the search using a much lower K.

14.3.11 K = 5

Top terms 01	Top terms 02	Top terms 03	Top terms 04	Top terms 05
governo	grazie	#coronavirus	grazie	grazie
presidente	lavoro	governo	pass	governo
paese	anni	#covid19	anni	anni
italiani	draghi	grazie	via	salvini
lavoro	#draghi	decreto	green	lavoro
grazie	grande	imprese	lavoro	#iostoconsalvini
anni	vaccini	emergenza	draghi	forza
politica	forza	l'italia	roma	grande
grande	vaccinale	misure	#greenpass	cittadini
via	lega	momento	grande	#lega

14.3.12

COMMENTHERE

14.3.13 Titles:

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
titles	1	2	3	4	5
	governo	grazie	#coronavirus	grazie	grazie
	presidente	lavoro	governo	pass	governo
	paese	anni	#covid19	anni	anni
	italiani	draghi	grazie	via	salvini
	lavoro	#draghi	decreto	green	lavoro
	grazie	grande	imprese	lavoro	#iostoconsalvini
	anni	vaccini	emergenza	draghi	forza
	politica	forza	l'italia	roma	grande
	grande	vaccinale	misure	#greenpass	cittadini
	via	lega	momento	grande	#lega

14.3.14

COMMENTHERE

14.4 Final try k = 22

14.4.1 Try with k = 22

14.4.2 Here i extract the most important terms from the model

Top terms 01	Top terms 02	Top terms 03	Top terms 04	Top terms 05	Top terms 06
vaccini	guerra	#draghi	grazie	#referendumgiustizia	#tokyo2020
aprile	#ucraina	draghi	mondo	luglio	#afghanistan
vaccinale	ucraina	governo	insieme	giustizia	talebani
pandemia	pace	#governodraghi	presidente	#ddlzan	agosto
@fratelliditalia	putin	lavoro	lavoro	gazebo	afghanistan
buon	marzo	paese	solidarietà	#primalitalia	medaglia
draghi	russia	@fratelliditalia	donne	giugno	oro
#draghi		presidente	paese	riforma	@fattoquotidiano
covid	popolo	buon	grande	firme	kabul
@pdnetwork	ucraino	vaccini	nazionale	#euro2020	medaglie

Top terms 7	Top terms 8	Top terms 9	Top terms 10	Top terms 11	Top terms 12
#coronavirus	maggio	sindaco	#iostoconsalvini	#iostoconsalvini	conte
misure	#ddlzan	ottobre	#oggivotolega	luglio	governo
momento	coprifuoco	pass	salvini	paese	crisi
coronavirus	@stampasgarbi	roma	#borgonzonipresidente	#salvini	#crisidigoverno
emergenza	vaccinale	@forza_italia	governo	giugno	#conte
grazie	#fratelliditalia	green	#salvini	#conte	maggioranza
l'emergenza	lavoro	città	#prescrizione	rilancio	paese
#covid19	#meloni	candidato	@fratelliditaiia	conte	presidente
casa	#coprifuoco	#roma	#m5s	#2giugno	gennaio
medici	sinistra	elettorale	#emiliaromagna	#recoveryfund	fiducia

Top terms 13	Top terms 14	Top terms 15	Top terms 16	Top terms 17	Top terms 18
#dpcm	#fase2	presidente	draghi	the	natale
#iostoconsalvini	maggio	#quirinale	pass	@fratelliditalia	dicembre
ottobre	#mes	repubblica	green	to	bilancio
#mes	lavoro	#presidentedellarepubblica	covid	of	anno
jole	ripartire	#mattarella	via	and	#atreju21
covid	aprile	#quirinale2022	#greenpass	bilancio	@fattoquotidiano
contagi	imprese	quirinale	@fratelliditalia	covid	#natale
@fratelliditalia	liquidità	gennaio	vaccinati	#mes	euro
#conte	#forzalombardia	mattarella	@stampasgarbi	natale	buon
de	#recoveryfund	david	pandemia	#covid	auguri

Top terms 19	Top terms 20	Top terms 21	Top terms 22
governo	settembre	novembre	grazie
italiani	#iovotono	de	lavoro
#covid19	elettorale	violenza	anni
paese	@fratelliditalia	pass	politica
#governo	#referendum	@fattoquotidiano	governo
imprese	scuola	donne	grande
conte	parlamentari	reddito	bene
l'italia	#iostoconsalvini	et	l'italia
crisi	referendum	@theskeptical_	forza
decreto	#processateancheme	renzi	via

14.4.3 Report on the analysis made with FER Puthon package

The package use the FER-2013 dataset created by Pierre Luc Carrier and Aaron Courville.

The dataset was created using the Google image search API to search for images of faces that match a set of 184 emotion-related keywords like “blissful”, “enraged,” etc. These keywords were combined with words related to gender, age or ethnicity, to obtain nearly 600 strings which were used as facial image search queries. The first 1000 images returned for each query were kept for the next stage of processing. OpenCV face recognition was used to obtain bounding boxes around each face in the collected images. Human labelers than rejected incorrectly labeled images, corrected the cropping if necessary, and filtered out some duplicate images. Approved, cropped images were then resized to 48x48 pixels and converted to grayscale. Mehdi Mirza and Ian Goodfellow prepared a subset of the images for this contest,and mapped the fine-grained emotion keywords into the same seven broad categories used in the Toronto Face Database [Joshua Susskind, Adam Anderson, and Geoffrey E. Hinton. The Toronto face dataset. Technical Report UTML TR 2010-001, U. Toronto, 2010.]. The resulting dataset contains 35887 images, with 4953 “Anger” images, 547 “Disgust” images, 5121 “Fear” images, 8989 “Happiness” images, 6077 “Sadness” images, 4002 “Surprise” images, and 6198 “Neutral” images. FER-2013 could theoretical suffer from label errors due to the way it was collected, but Ian Goodfellow found that human accuracy on FER-2013 was 65±5%.

66% ACCURACY REPORTED BY OCTAVIO ARRIAGA, Matias Valdenegro-Toro, Paul Plöger (Real-time Convolutional Neural Networks for Emotion and Gender Classification)