

notebook-only-with-manifestos

October 6, 2020

```
[1]: from config import *
     from utils import *
     import pandas as pd
```

1 Leer pickle

```
[2]: df = pd.read_pickle(pkl_file)
     df.head()
```

```
[2]:
```

	tweet_id	tweet_date	tweet_screen_name	\
0	1096977737930158080	2019-02-17	gabrielguerronv	
1	1096977884848082944	2019-02-17	jhonav00	
2	1096981630881882113	2019-02-17	Cristhyn3000	
3	1096982892759130112	2019-02-17	PelotonFusilero	
4	1096983080072556544	2019-02-17	carosu_92	

	tweet_text
0	@AlexTonello5RC @CompromisoRC5 @PaolaPabonC @M...
1	@PaolaPabonC @CompromisoRC5 @MashiRafael @Alex...
2	Que importante ha sido sentir el calor de los ...
3	Mi campaña no tiene pactos políticos ni aporte...
4	¡Lo más avergonzarte que he visto cómo mujer! ...

2 Top 10 candidatos más activos

```
[3]: candidates_tweets, candidates_lastnames = top_candidates(df,
    ↪only_manifestos=True)
```

Total tweets: 964439

Total tweets from candidates: 3345

The top 10 most active candidates are:

tweet_screen_name	
abenavidesgol1	1070
LuisaMaldonadoM	574
LoroHomero	509
PaolaVintimilla	279

```

mariasolcorral      201
CesarMontufar51     154
PacoMoncayo         150
EdgarJacomeT        140
juancaholguin       124
SevilaCarlos        91
Name: count, dtype: int64

```

3 Top 10 palabras por candidato

```

[4]: candidates_words, tf_idf, vocabulary, top_words_set = ↳
↳top_words(candidates_tweets)

```

The entire vocabulary has size 6908

Top 10 words for candidates:

```

LuisaMaldonadoM ['vamos', 'ciudad', 'gracias', 'pueblo', 'recuperar', 'hoy',
'lista', 'alcaldía', 'vecinos', 'quito']

```

```

LoroHomero      ['alcalde', 'jorge', 'yunda', 'quito', 'gracias', 'ciudad',
'hoy', 'trabajo', 'campaña', 'sur']

```

```

CesarMontufar51 ['quito', 'ciudad', 'gracias', 'apoyo', 'cada', 'recuperar',
'plan', 'quiteños', 'alcaldía', 'hoy']

```

```

abenavidesgol1 ['triunfo', 'copa', 'liga', 'empate', 'quito', 'ecuador',
'casa', 'ligapro', 'independiente', 'encuentro']

```

```

mariasolcorral  ['ciudad', 'plan', 'alcaldía', 'quito', 'mujeres', 'mañana',
'buena', 'día', 'propongo', 'público']

```

```

PaolaVintimilla ['quito', 'vamos', 'alcaldía', 'cambiar', 'forma', 'cambio',
'ciudad', 'aquí', 'debe', 'conocer']

```

```

EdgarJacomeT    ['marzo', 'plan', 'ciudad', 'propuestas', 'alcaldía', 'dial',
'proyecto', 'administración', 'crear', 'trabajo']

```

```

PacoMoncayo     ['juntos', 'ciudad', 'hacer', 'quito', 'quiteños', 'sistema',
'quiero', 'transporte', 'alcalde', 'gente']

```

```

SevilaCarlos    ['ciudad', 'juntos', 'construir', 'quito', 'excelente', 'vamos',
'oportunidades', 'propuestas', 'gracias', 'radio']

```

```

juancaholguin   ['gracias', 'quito', 'hola', 'hacer', 'puede', 'campaña',
'ciudad', 'cambiar', 'plan', 'buenos']

```

4 TF-IDF de una palabra

```
[5]: tf_idf[tf_idf['word'] == 'quito']
```

```
[5]:
```

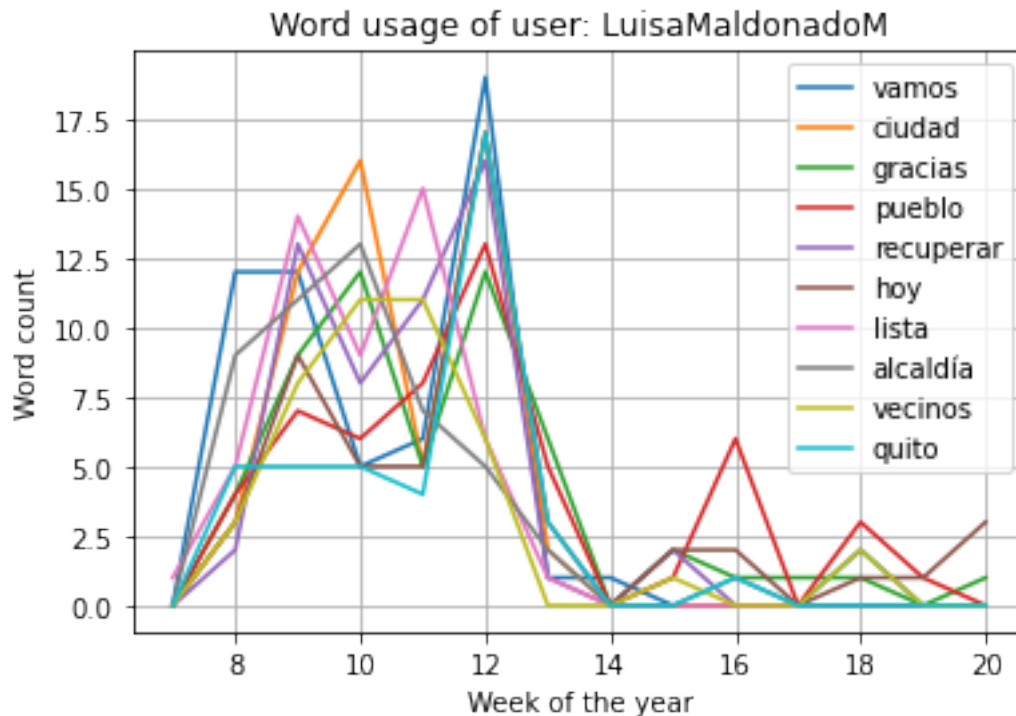
	word	tweet_id	tf	df	idf	tf-idf
181060	quito	1104738657435598848	0	40	4.410371	0.0
181061	quito	1098328701287256064	0	40	4.410371	0.0
181062	quito	1103744663817805824	0	40	4.410371	0.0
181063	quito	1104436152302534657	0	40	4.410371	0.0
181064	quito	1102455671138734080	0	40	4.410371	0.0
181065	quito	1127336105085472768	0	40	4.410371	0.0
181066	quito	1116563053011595264	0	40	4.410371	0.0
181067	quito	1115475660116512768	0	40	4.410371	0.0
181068	quito	1122244982696488968	0	40	4.410371	0.0
181069	quito	1109546009745592320	0	40	4.410371	0.0
181070	quito	1122358603866755073	0	40	4.410371	0.0
181071	quito	1109625610123395074	0	40	4.410371	0.0
181072	quito	1106965097921105922	0	40	4.410371	0.0
181073	quito	1097313112053698560	0	40	4.410371	0.0
181074	quito	1100005539013103616	0	40	4.410371	0.0
181075	quito	1117185566926430214	0	40	4.410371	0.0
181076	quito	1116907640003682304	0	40	4.410371	0.0
181077	quito	1098651932971937792	0	40	4.410371	0.0
181078	quito	1100579663670398977	0	40	4.410371	0.0
181079	quito	1107976532981698560	0	40	4.410371	0.0
181080	quito	1105242921228533760	0	40	4.410371	0.0
181081	quito	1126701310420574208	0	40	4.410371	0.0
181082	quito	1107821056004378624	0	40	4.410371	0.0
181083	quito	1109099685246054400	0	40	4.410371	0.0
181084	quito	1109825562363744258	0	40	4.410371	0.0
181085	quito	1126637322643095552	0	40	4.410371	0.0
181086	quito	1103039804906844160	0	40	4.410371	0.0
181087	quito	1117569121099636742	0	40	4.410371	0.0
181088	quito	1115019020678332422	0	40	4.410371	0.0
181089	quito	1104104117729968129	0	40	4.410371	0.0
...
184322	quito	1112691898320338950	0	40	4.410371	0.0
184323	quito	1104164154733223936	0	40	4.410371	0.0
184324	quito	1109202158237433861	0	40	4.410371	0.0
184325	quito	1106356782199382018	0	40	4.410371	0.0
184326	quito	1100768756530335744	0	40	4.410371	0.0
184327	quito	1122597033054035968	0	40	4.410371	0.0
184328	quito	1107644053313736704	0	40	4.410371	0.0
184329	quito	1113983340196110336	0	40	4.410371	0.0
184330	quito	1127406190047133696	0	40	4.410371	0.0
184331	quito	1104553717385318430	0	40	4.410371	0.0
184332	quito	1098341618699055104	0	40	4.410371	0.0

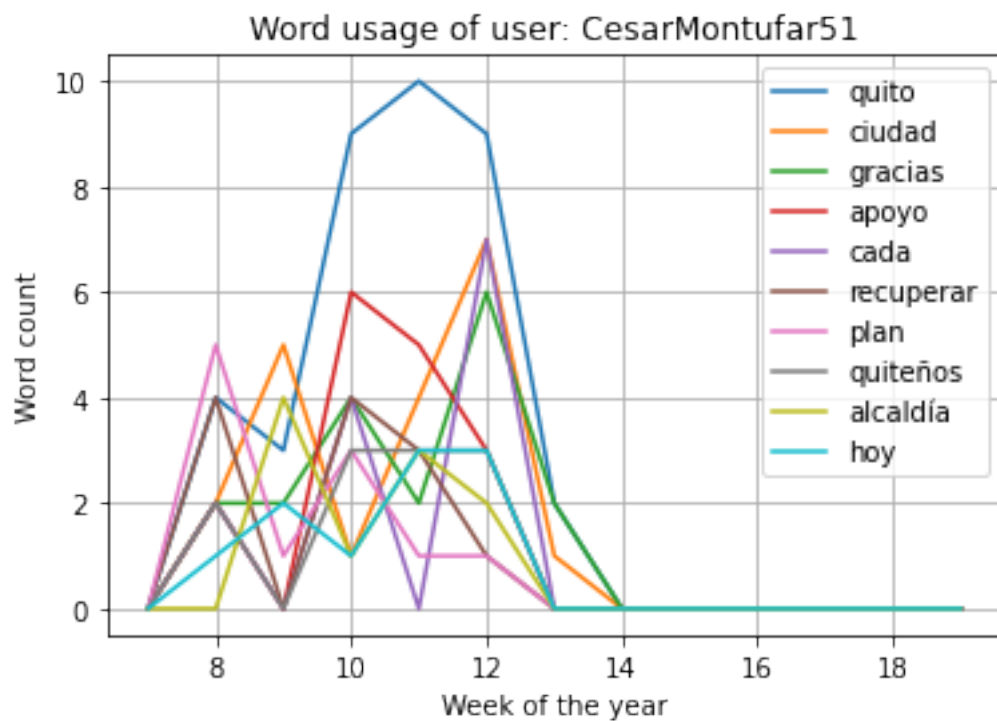
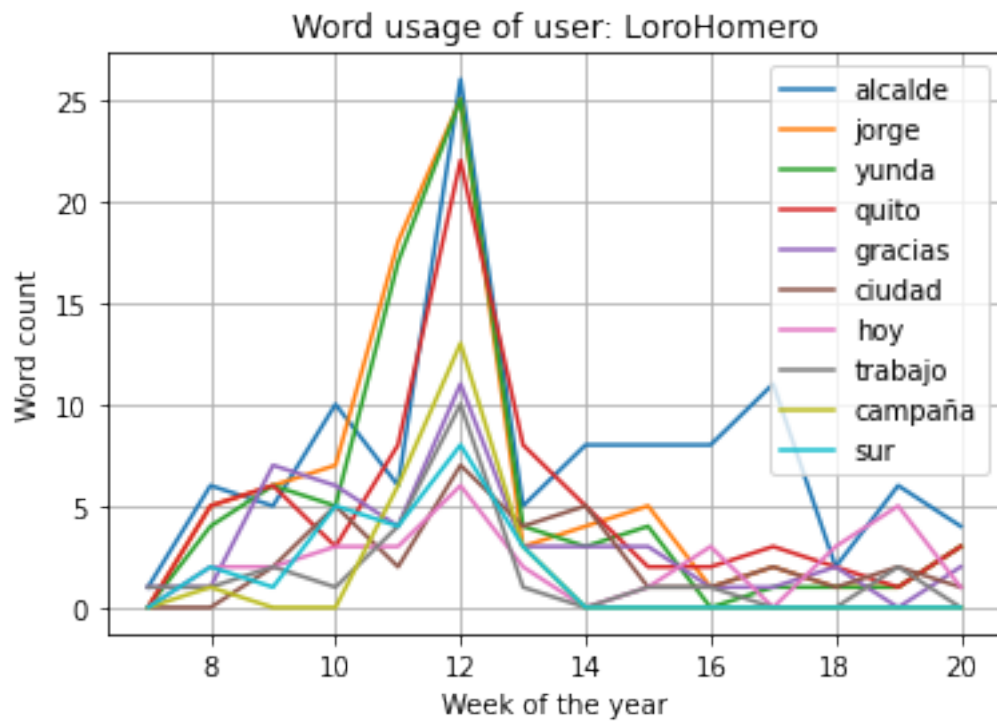
184333	quito	1098042233071325184	0	40	4.410371	0.0
184334	quito	1107288058003505154	0	40	4.410371	0.0
184335	quito	1105452723124822018	0	40	4.410371	0.0
184336	quito	1118152481803055104	0	40	4.410371	0.0
184337	quito	1117074660854763521	0	40	4.410371	0.0
184338	quito	1120893498209783809	0	40	4.410371	0.0
184339	quito	1109825583276544000	0	40	4.410371	0.0
184340	quito	1107276718551851008	0	40	4.410371	0.0
184341	quito	1126326288065748992	0	40	4.410371	0.0
184342	quito	1123676464489074690	0	40	4.410371	0.0
184343	quito	1123763797305503750	0	40	4.410371	0.0
184344	quito	1109187723796115462	0	40	4.410371	0.0
184345	quito	1104781564930596864	0	40	4.410371	0.0
184346	quito	1106057666982379520	0	40	4.410371	0.0
184347	quito	1123246172750123008	0	40	4.410371	0.0
184348	quito	1112195154394386433	0	40	4.410371	0.0
184349	quito	1099053303751761921	0	40	4.410371	0.0
184350	quito	1115466340985384960	0	40	4.410371	0.0
184351	quito	1109825037937373184	0	40	4.410371	0.0

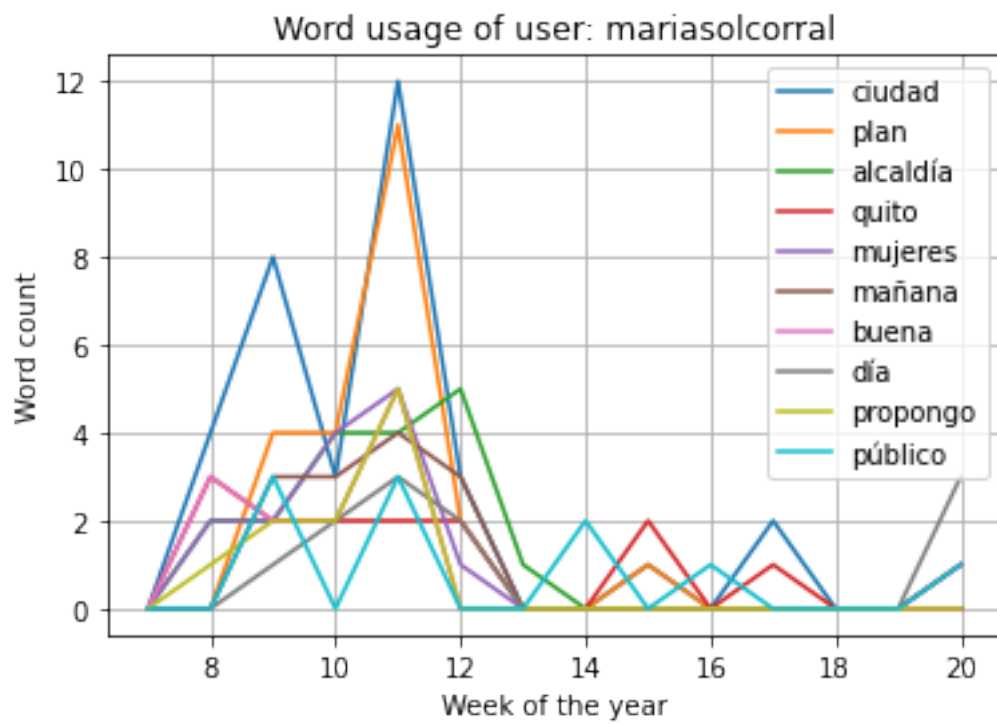
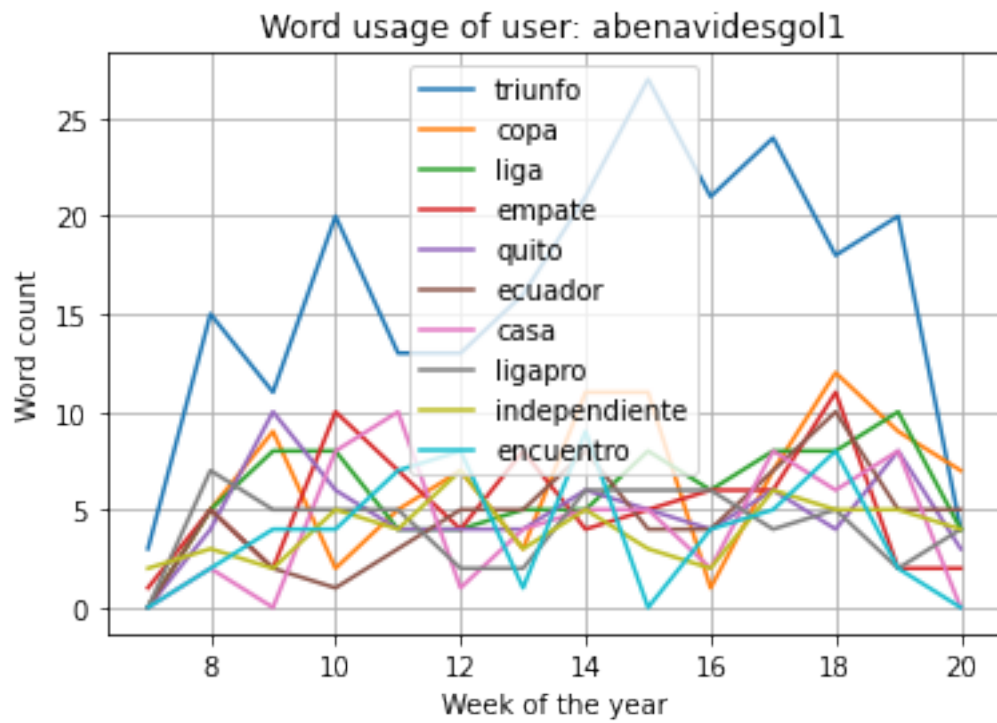
[3292 rows x 6 columns]

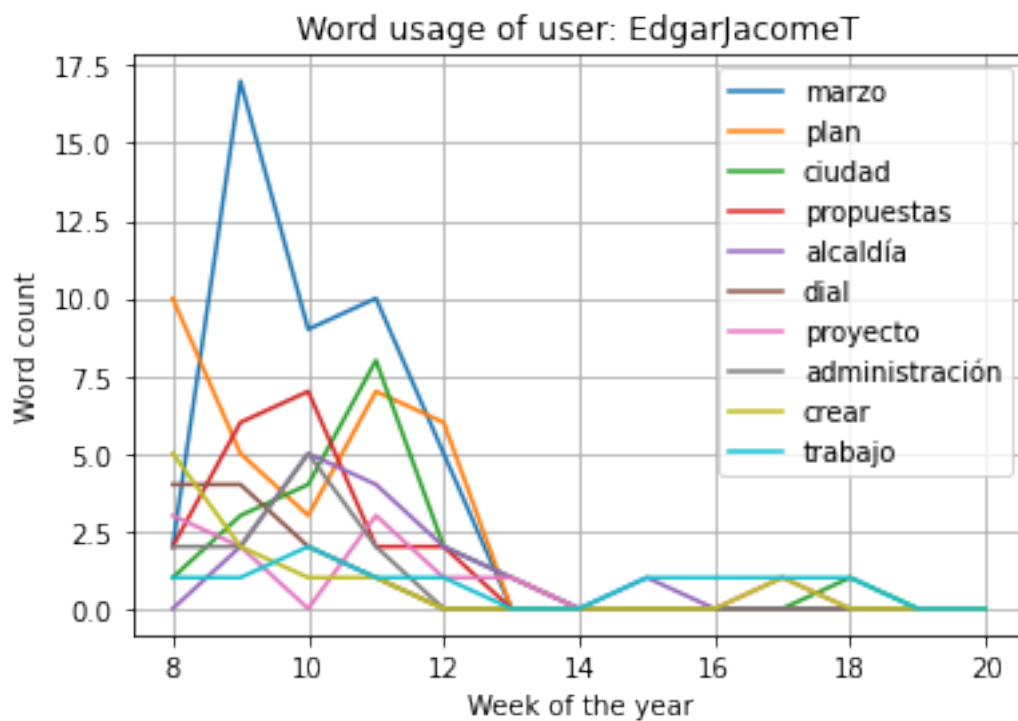
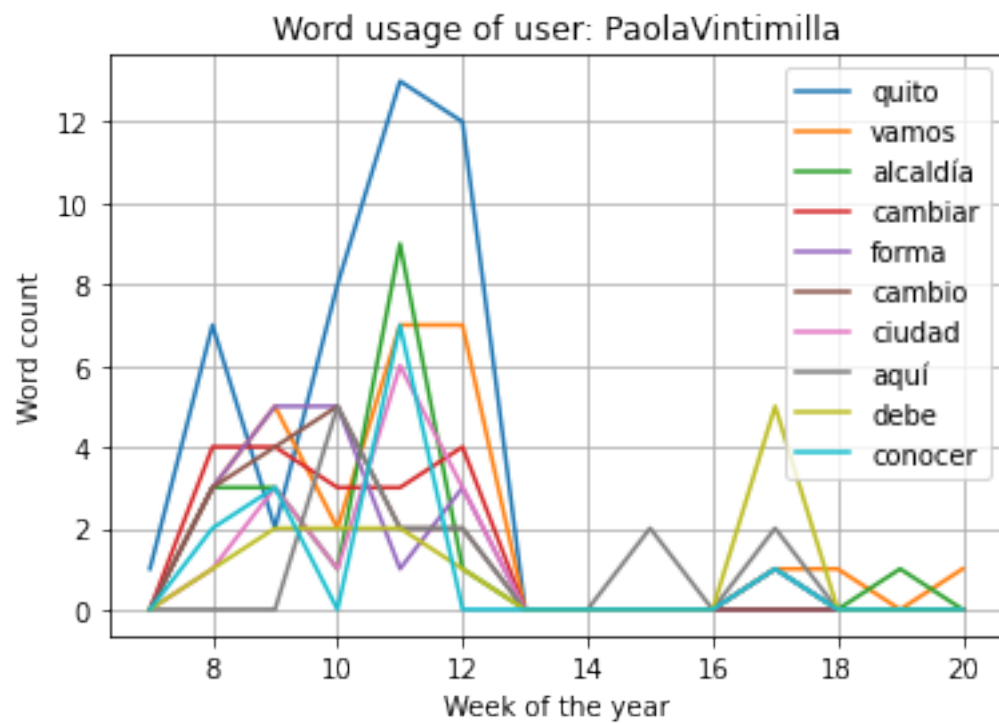
5 Timeline de palabras por candidato

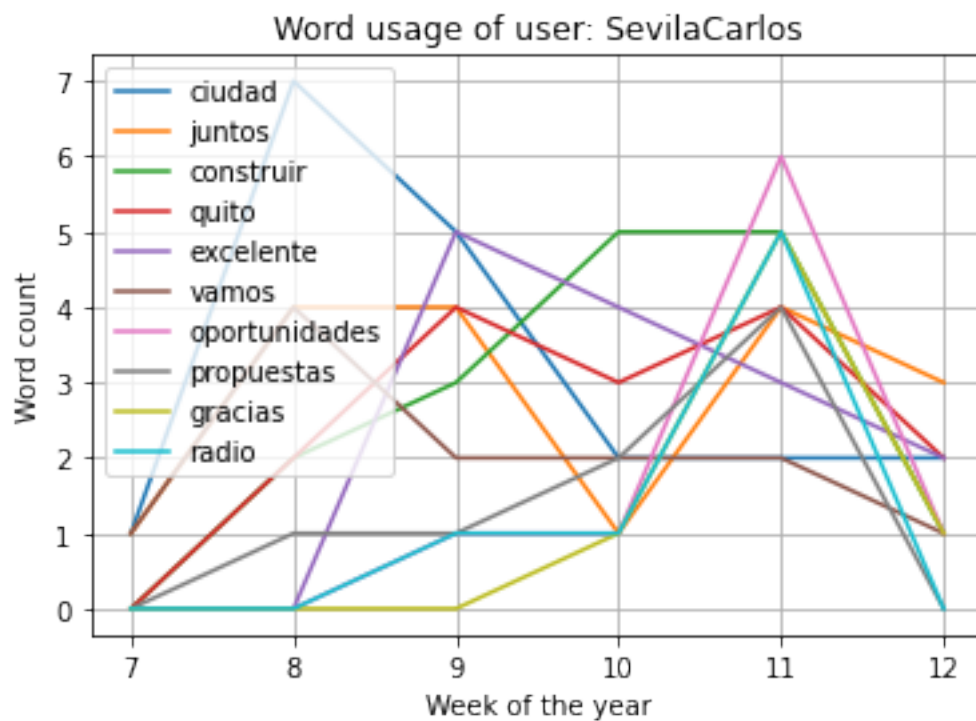
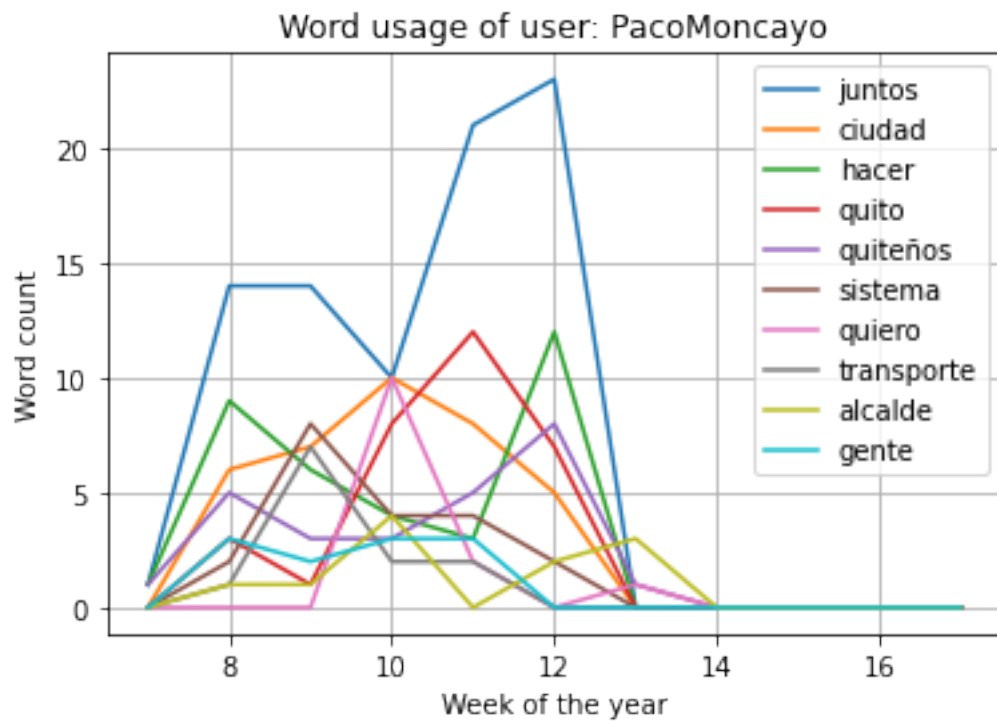
```
[6]: time_line(candidates_words, candidates_tweets)
```

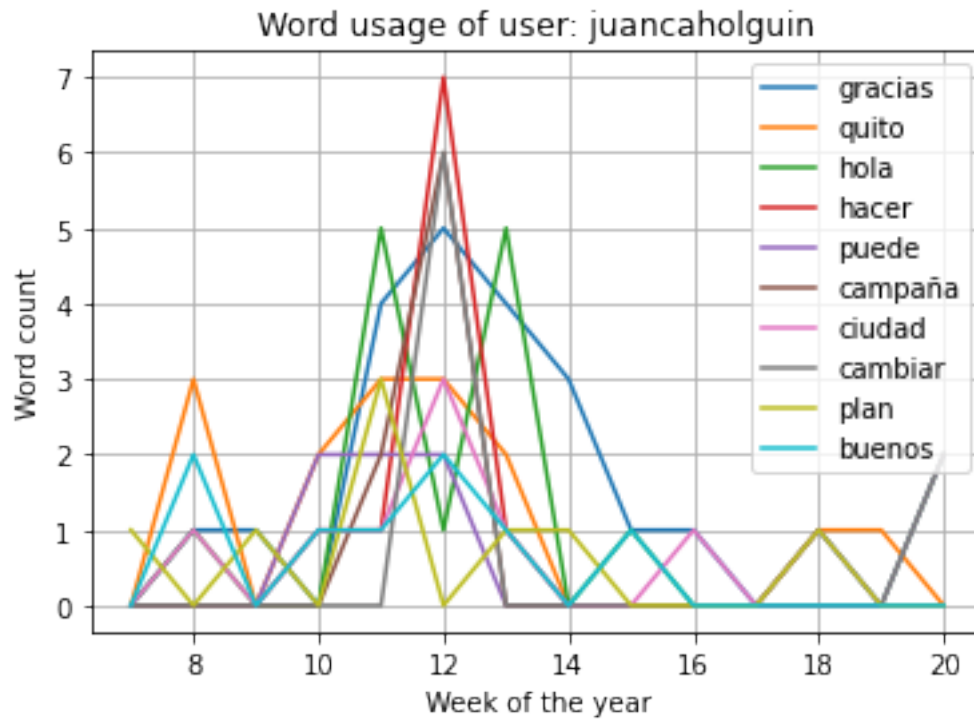












6 Ranking de candidatos por query

```
[7]: total_user_vectors = ranking('alcaldía quito ciudad trabajamos empresa',  
    ↪ vocabulary, candidates_tweets, tf_idf)
```

The query vector is:

[illegible]

[illegible]

[illegible]


```
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
```

```
2.      LoroHomero
Total score:      13.478511224672474
Top ranked tweet:
En el aeropuerto de Miami, hay una aérea para las mascotas que están de viaje,
que me pareció genial podrían implementar en quito señor Alcalde @LoroHomero
https://t.co/7NNAePoYjV
Tweet score: 0.7217611853520829
```

```
4.      EdgarJacomeT
Total score:      8.714136432638357
Top ranked tweet:
@LiberaEcuador @MANimalistas | La experiencia política no es un condicionante
para aspirar a la alcaldía de #Quito. Por culpa de los expertos políticos
nuestra ciudad está sumida en el caos y corrupción. Mis capacidades y
experiencia me hacen merecedor de esta candidatura. https://t.co/6FN4DzASCz
Tweet score: 0.6921421756540751
```

6. PaolaVintimilla
Total score: 6.392490065720225

Top ranked tweet:

"El Quito Vivo que queremos será respaldado por la Policía Nacional para proteger a los ciudadanos" @PaolaVintimilla #alcalde #quito #proteger #vida #nocturna #24 #horas #ecuador #democraciatv <https://t.co/iBAuNJbYuF>

Tweet score: 0.7217611853520829

7. CesarMontufar51

Total score: 4.835486828025434

Top ranked tweet:

@GuayaquilQ, @xbuendia2019, @mariasolcorral, @PabloDavalos63, @CesarMontufar51, @AndresPasquel8, @EdgarJacomeT, @victorhugoerazo y @JCSolines en el diálogo con los candidatos a la alcaldía de #Quito <https://t.co/r2MqAhUcos>

Tweet score: 0.6062940713633117

8. SevillaCarlos

Total score: 4.222600641643242

Top ranked tweet:

#ConociendoAlCandidato #AlcaldíaQuito

Conozcamos el perfil del candidato a la alcaldía por el movimiento Libertad Es Pueblo, @SevilaCarlos

#Decisión2019 <https://t.co/xT5GueyBaK>

Tweet score: 0.6062940713633117

9. abenavidesgol1

Total score: 2.066218257332785

Top ranked tweet:

#QuéProponen

El candidato a la alcaldía por la Lista 3, @abenavidesgol1, plantea convertir a Quito como a una ciudad digital y además la creación de paradas inteligentes para el Metro de Quito.

(Foto de: @elcomerciocom) <https://t.co/HCPgDifTiD>

Tweet score: 0.6921421756540751

10. juancaholguin

Total score: 1.612842586319998

Top ranked tweet:

El empleo es el eje principal de nuestra propuesta. Tenemos que generar empleo para que esta ciudad crezca. Proponemos ordenar las ventas creando corredores comerciales donde los emprendedores urbanos puedan trabajar.

<https://t.co/Z1CtEAp2T1>

Tweet score: 0.3338686723082233

7 Similaridad de actividad entre candidatos

```
[8]: similarity(total_user_vectors)
```

```
[8]:          LuisaMaldonadoM LoroHomero CesarMontufar51 abenavidesgol1 \
LuisaMaldonadoM          1      0.621933      0.557325      0.0466849
LoroHomero              0.621933          1      0.501839      0.0653064
CesarMontufar51        0.557325      0.501839          1      0.0180435
abenavidesgol1         0.0466849  0.0653064      0.0180435          1
mariasolcorral         0.429554      0.467584      0.623108      0.0530222
PaolaVintimilla        0.489592      0.478597      0.479968      0.0564382
EdgarJacomeT           0.39798      0.544963      0.441653      0.028618
PacoMoncayo            0.321316      0.381728      0.405853      0.0148381
SevilaCarlos           0.381675      0.358113      0.445644      0.0408469
juancaholguin          0.430351      0.564235      0.347314      0.0232574

          mariasolcorral PaolaVintimilla EdgarJacomeT PacoMoncayo \
LuisaMaldonadoM        0.429554          0.489592      0.39798      0.321316
LoroHomero             0.467584          0.478597      0.544963      0.381728
CesarMontufar51        0.623108          0.479968      0.441653      0.405853
abenavidesgol1         0.0530222      0.0564382      0.028618      0.0148381
mariasolcorral         1              0.359115      0.566367      0.3644
PaolaVintimilla        0.359115          1      0.469465      0.293748
EdgarJacomeT           0.566367          0.469465          1      0.234658
PacoMoncayo            0.3644          0.293748      0.234658          1
SevilaCarlos           0.348307          0.37861      0.36154      0.459262
juancaholguin          0.421804          0.479473      0.528995      0.326087

          SevilaCarlos juancaholguin
LuisaMaldonadoM        0.381675      0.430351
LoroHomero             0.358113      0.564235
CesarMontufar51        0.445644      0.347314
abenavidesgol1         0.0408469      0.0232574
mariasolcorral         0.348307      0.421804
PaolaVintimilla        0.37861      0.479473
EdgarJacomeT           0.36154      0.528995
PacoMoncayo            0.459262      0.326087
SevilaCarlos           1              0.222687
juancaholguin          0.222687          1
```

8 Similaridad entre actividad de candidatos con sus manifiestos

```
[9]: deviation(candidates_words, candidates_lastnames, vocabulary, ↵
↵candidates_tweets, total_user_vectors, tf_idf)
```

Euclidean distance of reduced vocabulary vectors of tweets vs. manifestos

