Palabra más repetida

```
clear; clc; warning off;
[x,t]=xlsread('DBTweets.xlsx');
t=t(2:end); tc=string(t);
tc=unique(tc); tc=eraseURLs(tc);
CD=tokenizedDocument(tc);
CD=normalizeWords(CD, "Style", "lemma");
CD=erasePunctuation(CD);
```

Parte 1

```
allwords=CD(1).Vocabulary;
for i=2:1:length(CD)
    allwords=cat(2,allwords,CD(i).Vocabulary);
end
[C,ia,ic]=unique(allwords);
a=accumarray(ic,1); id=find(a==max(a));
fprintf('La palabra más repetida es "%s" y se repite %d veces \n',C(id),max(a));
La palabra más repetida es "de" y se repite 667 veces
```

```
wordcloud(allwords);
```

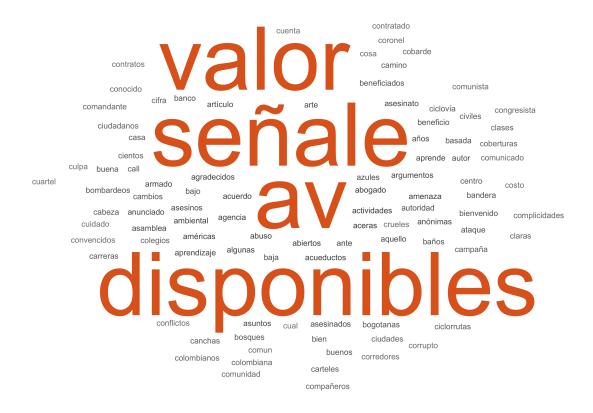
```
derechos
                         desarrollo
                              sociales
                                         #medellínsetransformó
                              gobierno
                                                estado
              presidente
                                                    gracias
                                niños
                                                             @ivanduque
                                   colombia esta
             sobre
   política
                                                               general
                                                 hacer x001a
       ya <sub>ha</sub>
               vida
                                                             nuevo seguridad
               mi
      años
hace
                 líderes
                                       paz
                                                               ciudad
quien
                                             toda
 jóvenes
                                                                     hay
 ahora
                                                           este
                              eso
                                                                     gran
   mejor
                                                          todos proyecto
  pública
          parque
                                                       hoy bogotá
  público
         nuestra
                                      país
         millones
                                                       x001ax001a
                             debe
                                                          persona
                                        colombianos
                @ficogutierrez contra
                                        porque
                                    @antanasmockus
```

```
cn=wordCloudCounts(allwords);
head(cn)
```

ans = 8×2 table		
	Word	Count
1	"de"	667
2	"la"	605
3	"en"	493
4	"у"	450
5	"el"	409
6	"que"	376
7	"los"	375
8	"del"	235

Parte 2

```
CB=bagOfWords(CD);
aw=CB.Vocabulary;
wordcloud(aw);
```



Parte 3

