

# Gráfico de nube

Etna Carolina Cortés Martínez

23/11/2021

## Gráfico de nube del cuento “Los tres cerditos”

### Introducción

Una nube de palabras (**también conocida como nube de etiquetas o de texto**) , es una representación visual sobre la regularidad en la que se repiten las palabras, es utilizada principalmente para resaltar la manera en la que se repiten ciertos términos, cuantas más veces una palabra clave está presente en un conjunto de datos, esa palabra aparecerá más grande y más remarcada.

### RESUMEN:

El presente proyecto contiene el tratamiento de un texto literario dirigido a niños pequeños llamado “**Los tres cerditos**”.

### DATOS:

*La siguiente información se obtuvo de:*

**Artmann, P. (2020). Cuento de los tres cerditos. juegos educativos para niños. Recuperado 23 de noviembre de 2021, de <https://arbolabc.com/cuentos-clasicos-infantiles/los-tres-cochinitos>**

### 1.- Se instalo con anteriridad los presentes paquetes.

```
install.packages("tm")  
install.packages("NLP")  
install.packages("SnowballC")  
install.packages("wordcloud")  
install.packages("RColorBrewer")
```

## 2.- cargamos las librerias correspondientes

```
library(tm)
```

```
library(NLP)
```

```
library(SnowballC)
```

```
library(RColorBrewer)
```

```
library(wordcloud)
```

## CUENTO " Los tres cerditos"

NOTA: nuestro texto lo guardamos en bloc de notas en formato txt.

```
texto <- readLines("Cuento.txt")
```

```
texto = iconv(texto, to="ASCII//TRANSLIT")
```

```
texto = Corpus(VectorSource(texto))
```

## LIMPIAMOS NUESTRO TEXTO CON EL COMANDO tm\_map

1.- Convertir todos los datos en minuscula (A=a)

```
discurso=tm_map(texto,tolower)
```

2.- Quitamos los espacios en blanco

```
discurso =tm_map(discurso, stripWhitespace)
```

3.- Quitamos la puntuacion

```
discurso = tm_map(discurso, removePunctuation)
```

4.- Quitamos los numeros.

```
discurso = tm_map(discurso, removeNumbers)
```

5.- Mostramos palabras vacias y genericas.

```
stopwords("spanish")
```

|    |       |             |              |                |               |              |
|----|-------|-------------|--------------|----------------|---------------|--------------|
| ## | [1]   | "de"        | "la"         | "que"          | "el"          | "en"         |
| ## | [6]   | "y"         | "a"          | "los"          | "del"         | "se"         |
| ## | [11]  | "las"       | "por"        | "un"           | "para"        | "con"        |
| ## | [16]  | "no"        | "una"        | "su"           | "al"          | "lo"         |
| ## | [21]  | "como"      | "más"        | "pero"         | "sus"         | "le"         |
| ## | [26]  | "ya"        | "o"          | "este"         | "sí"          | "porque"     |
| ## | [31]  | "esta"      | "entre"      | "cuando"       | "muy"         | "sin"        |
| ## | [36]  | "sobre"     | "también"    | "me"           | "hasta"       | "hay"        |
| ## | [41]  | "donde"     | "quien"      | "desde"        | "todo"        | "nos"        |
| ## | [46]  | "durante"   | "todos"      | "uno"          | "les"         | "ni"         |
| ## | [51]  | "contra"    | "otros"      | "ese"          | "eso"         | "ante"       |
| ## | [56]  | "ellos"     | "e"          | "esto"         | "mí"          | "antes"      |
| ## | [61]  | "algunos"   | "qué"        | "unos"         | "yo"          | "otro"       |
| ## | [66]  | "otras"     | "otra"       | "él"           | "tanto"       | "esa"        |
| ## | [71]  | "estos"     | "mucho"      | "quienes"      | "nada"        | "muchos"     |
| ## | [76]  | "cual"      | "poco"       | "ella"         | "estar"       | "estas"      |
| ## | [81]  | "algunas"   | "algo"       | "nosotros"     | "mi"          | "mis"        |
| ## | [86]  | "tú"        | "te"         | "ti"           | "tu"          | "tus"        |
| ## | [91]  | "ellas"     | "nosotras"   | "vosotros"     | "vosotras"    | "os"         |
| ## | [96]  | "mío"       | "mía"        | "míos"         | "mías"        | "tuyo"       |
| ## | [101] | "tuya"      | "tuyos"      | "tuyas"        | "suyo"        | "suya"       |
| ## | [106] | "suyos"     | "suyas"      | "nuestro"      | "nuestra"     | "nuestros"   |
| ## | [111] | "nuestras"  | "vuestro"    | "vuestra"      | "vuestros"    | "vuestras"   |
| ## | [116] | "esos"      | "esas"       | "estoy"        | "estás"       | "está"       |
| ## | [121] | "estamos"   | "estáis"     | "están"        | "esté"        | "estés"      |
| ## | [126] | "estemos"   | "estéis"     | "estén"        | "estaré"      | "estarás"    |
| ## | [131] | "estará"    | "estaremos"  | "estaréis"     | "estarán"     | "estaría"    |
| ## | [136] | "estarías"  | "estaríamos" | "estaríais"    | "estarían"    | "estaba"     |
| ## | [141] | "estabas"   | "estábamos"  | "estabais"     | "estaban"     | "estuve"     |
| ## | [146] | "estuviste" | "estuvo"     | "estuvimos"    | "estuvisteis" | "estuvieron" |
| ## | [151] | "estuviera" | "estuvieras" | "estuviéramos" | "estuvierais" | "estuvieran" |
| ## | [156] | "estuviese" | "estuvieses" | "estuviésemos" | "estuvieseis" | "estuviesen" |
| ## | [161] | "estando"   | "estado"     | "estada"       | "estados"     | "estadas"    |
| ## | [166] | "estad"     | "he"         | "has"          | "ha"          | "hemos"      |
| ## | [171] | "habéis"    | "han"        | "haya"         | "hayas"       | "hayamos"    |
| ## | [176] | "hayáis"    | "hayan"      | "habré"        | "habrás"      | "habrá"      |
| ## | [181] | "habremos"  | "habréis"    | "habrán"       | "habría"      | "habrías"    |
| ## | [186] | "habríamos" | "habríais"   | "habrían"      | "había"       | "habías"     |
| ## | [191] | "habíamos"  | "habíais"    | "habían"       | "hube"        | "hubiste"    |
| ## | [196] | "hubo"      | "hubimos"    | "hubisteis"    | "hubieron"    | "hubiera"    |
| ## | [201] | "hubieras"  | "hubiéramos" | "hubierais"    | "hubieran"    | "hubiese"    |
| ## | [206] | "hubieses"  | "hubiésemos" | "hubieseis"    | "hubiesen"    | "habiendo"   |
| ## | [211] | "habido"    | "habida"     | "habidos"      | "habidas"     | "soy"        |
| ## | [216] | "eres"      | "es"         | "somos"        | "sois"        | "son"        |
| ## | [221] | "sea"       | "seas"       | "seamos"       | "seáis"       | "sean"       |
| ## | [226] | "seré"      | "serás"      | "será"         | "seremos"     | "seréis"     |
| ## | [231] | "serán"     | "sería"      | "serías"       | "seríamos"    | "seríais"    |
| ## | [236] | "serían"    | "era"        | "eras"         | "éramos"      | "erais"      |
| ## | [241] | "eran"      | "fui"        | "fuiste"       | "fue"         | "fuimos"     |
| ## | [246] | "fuisteis"  | "fueron"     | "fuera"        | "fueras"      | "fuéramos"   |
| ## | [251] | "fuerais"   | "fueran"     | "fuese"        | "fueses"      | "fuésemos"   |
| ## | [256] | "fueseis"   | "fuesen"     | "siendo"       | "sido"        | "tengo"      |
| ## | [261] | "tienes"    | "tiene"      | "tenemos"      | "tenéis"      | "tienen"     |
| ## | [266] | "tenga"     | "tengas"     | "tengamos"     | "tengáis"     | "tengan"     |

```
## [271] "tendré"      "tendrás"      "tendrá"      "tendremos"   "tendréis"
## [276] "tendrán"     "tendría"      "tendrías"    "tendríamos"  "tendríais"
## [281] "tendrían"    "tenía"        "tenías"      "teníamos"    "teníais"
## [286] "tenían"      "tuve"         "tuviste"     "tuvo"        "tuvimos"
## [291] "tuvisteis"   "tuvieron"     "tuviera"     "tuvieras"    "tuviéramos"
## [296] "tuvierais"   "tuvieran"     "tuviese"     "tuvieses"    "tuviésemos"
## [301] "tuvieseis"   "tuviesen"     "teniendo"    "tenido"      "tenida"
## [306] "tenidos"     "tenidas"      "tened"
```

6.-Quitamos palabras genericas.

```
discurso=tm_map(discurso, removeWords, stopwords("spanish"))
```

## DATA FRAME DE PALABRAS CON SU FRECUENCIA

1.- Creamos matriz de letras.

```
letras= TermDocumentMatrix(discurso)
```

```
findFreqTerms(letras, lowfreq=5)
```

```
## [1] "cerdita" "cerditos" "dijo"      "mama"      "tres"      "casita"
## [7] "cerdito" "medio"     "feroz"     "lobo"      "entrar"    "nunca"
## [13] "casa"
```

```
matrix=as.matrix(letras)
```

2.- Lo ordenamos y sumamos las letras de nuestra matriz.

```
vector <- sort(rowSums(matrix),decreasing=TRUE)
```

3.- Creamos la data con las palabras y su frecuencia.

```
dataletras <- data.frame(word= names(vector),frecuencia=vector)
```

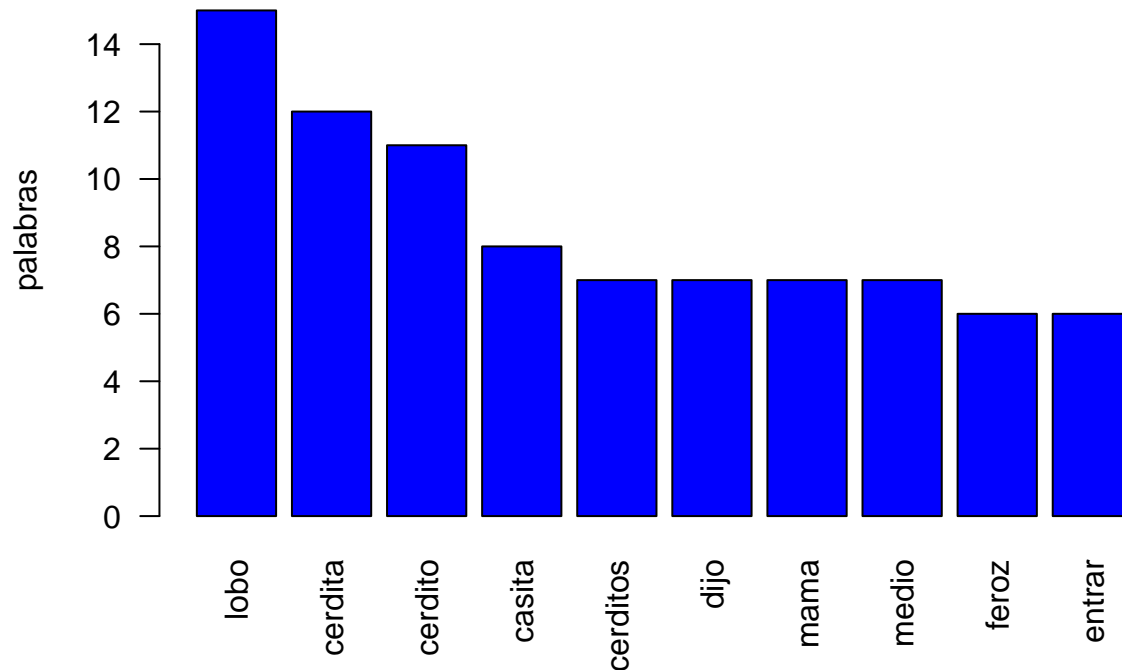
## METODOLOGÍA DE ANÁLISIS:

Realizamos un análisis de texto para poder visualizar la frecuencia con la que se repite cierto número de términos, esto a través de la limpieza de texto y creación de una matriz de letras, logrando graficar dichos datos de manera clara y precisa.

## GRAFICAMOS FRECUENCIA DE LAS PALABRAS

```
barplot(dataletras[1:10,]$freq, las = 2, names.arg = dataletras[1:10,]$word,
        col = "blue", main = "PALABRAS MAS FRECUENTES", ylab = "Frecuencia de
palabras")
```

## PALABRAS MAS FRECUENTES

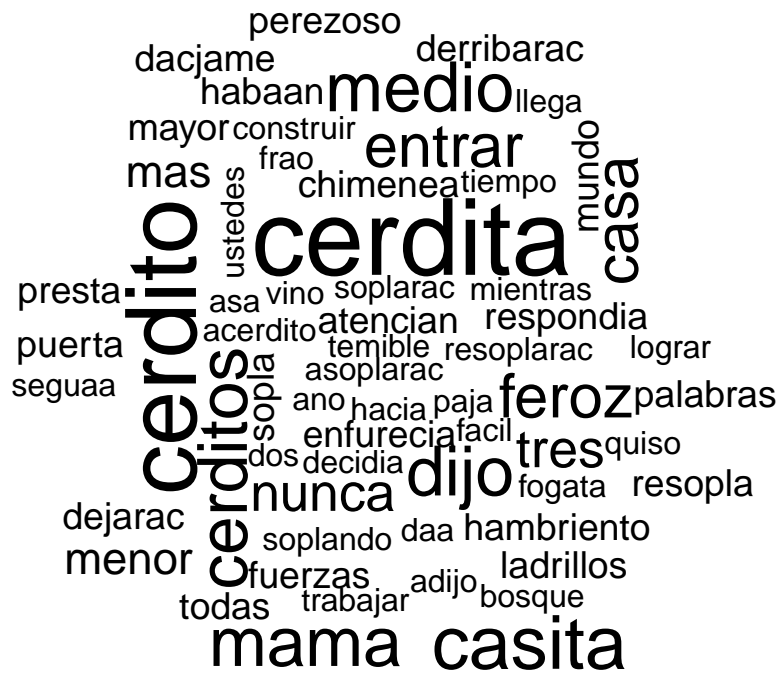


## RESULTADOS:

A través del tratamiento de los datos pudimos notar que la palabra con mayor frecuencia es **Lobo** con 14 veces, al contrario de las palabras **Feroz** y **Entrar** que solo se repiten en 7 ocasiones, dando como resultado una nube con la palabra **Lobo**, más grande que los otros términos.

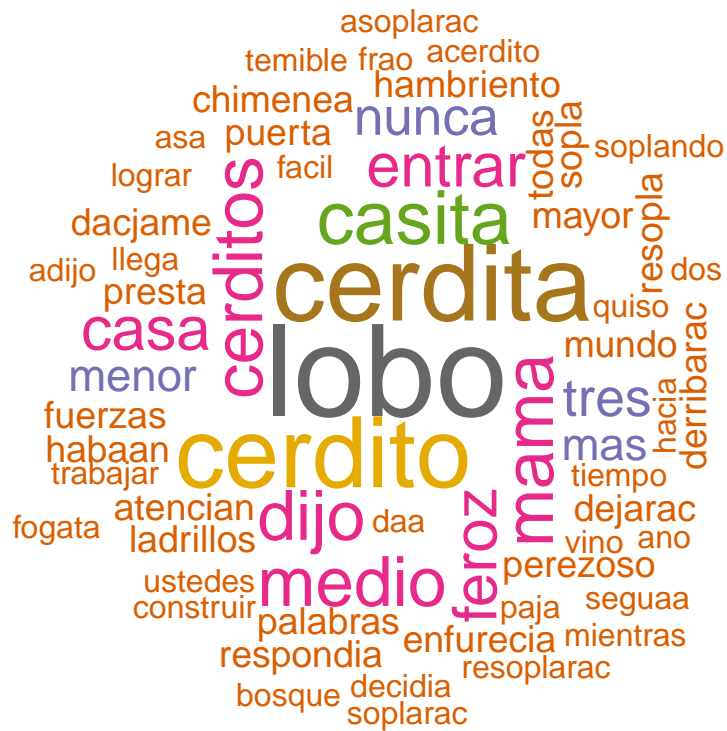
## GRAFICAMOS LA NUBE DE PALABRAS

```
wordcloud(words = dataletras$word, freq = dataletras$freq, min.freq = 2,  
          max.words=100)
```



1.- En el centro la palabra mas importante.

```
wordcloud(words = dataletras$word, freq = dataletras$freq, min.freq = 2,
max.words=100, random.order=FALSE, rot.per=0.2,
colors=brewer.pal(8, "Dark2"))
```



## # REFERENCIAS:

- 1.- Graficar NUBE de PALABRAS con R Studio. (2019, 30 diciembre). YouTube. Recuperado 23 de noviembre de 2021, de (<https://www.youtube.com/watch?v=tihnYJhZCHs>)
- 2.- Artmann, P. (2020). Cuento de los tres cerditos. juegos educativos para niños. Recuperado 23 de noviembre de 2021, de (<https://arbolabc.com/cuentos-clasicos-infantiles/los-tres-cochinitos>)