```
# Instalando quanteda package e relacionados, quer dizer estou basicamente replicando a
# partir do tutorial https://tutorials.quanteda.io/introduction/install/
# BASIC PACKAGES
install.packages("quanteda")
install.packages("quanteda.textmodels")
install.packages("quanteda.textstats")
install.packages("quanteda.textplots")
install.packages("readtext")
    Installing package into '/usr/local/lib/R/site-library'
     (as 'lib' is unspecified)
     Installing package into '/usr/local/lib/R/site-library'
     (as 'lib' is unspecified)
     also installing the dependency 'glmnet'
     Installing package into '/usr/local/lib/R/site-library'
     (as 'lib' is unspecified)
     also installing the dependency 'proxyC'
     Installing package into '/usr/local/lib/R/site-library'
     (as 'lib' is unspecified)
     also installing the dependencies 'coda', 'extrafontdb', 'Rttf2pt1', 'statnet.common', 'extrafont', 'ggrepel', 'sna', 'igraph', 'netw
     Installing package into '/usr/local/lib/R/site-library'
     (as 'lib' is unspecified)
     also installing the dependencies 'bitops', 'qpdf', 'RCurl', 'rjson', 'ndjson', 'antiword', 'pdftools', 'readODS', 'streamR', 'stripr
     Warning message in install.packages("readtext"):
     "installation of package 'pdftools' had non-zero exit status"
     Warning message in install.packages("readtext"):
     "installation of package 'readtext' had non-zero exit status"
#
# We will also use extra datasets in tutorials that are available in quanteda.corpora. This package
# is not on CRAN, but can be installed with the install_github() function from the devtools package
install.packages("devtools") # get devtools to install quanteda.corpora
devtools::install_github("quanteda/quanteda.corpora")
# A função a seguir atualiza os pacotes a fim evitar erros
update.packages()
Installing package into '/usr/local/lib/R/site-library'
     (as 'lib' is unspecified)
     also installing the dependencies 'httpuv', 'xtable', 'fontawesome', 'sourcetools', 'later', 'promises', 'textshaping', 'shiny', 'dow
     Warning message in install.packages("devtools"):
     "installation of package 'textshaping' had non-zero exit status"
     Warning message in install.packages("devtools"):
     "installation of package 'ragg' had non-zero exit status"
     Warning message in install.packages("devtools"):
     "installation of package 'pkgdown' had non-zero exit status" Warning message in install.packages("devtools"):
     "installation of package 'devtools' had non-zero exit status"
     Downloading GitHub repo quanteda/quanteda.corpora@HEAD
     * checking for file '/tmp/Rtmpgq2xzK/remotes9d753ccbef8/quanteda-quanteda.corpora-ec4b76d/DESCRIPTION' ... OK
     * preparing 'quanteda.corpora':
     * checking DESCRIPTION meta-information ... OK
       checking for LF line-endings in source and make files and shell scripts
     * checking for empty or unneeded directories
     Removed empty directory 'quanteda.corpora/sources'
     * building 'quanteda.corpora_0.9.2.tar.gz'
     Installing package into '/usr/local/lib/R/site-library'
     (as 'lib' is unspecified)
    4
```

```
#------
# ADDITIONAL PACKAGES
# The tutorials do not cover syntactical analysis, but you should install spacyr for part-of-speech
# tagging, entity recognition, and dependency parsing. It provides an interface to the spaCy
# library and works well with quanteda. To install it we have to have Python installed to use the
# spacyr package. For more info check: https://github.com/quanteda/spacyr/blob/master/README.md
install.packages("spacyr")
# Use newsmap to classify documents based on "seed words" in dictionaries and the seededlda package to run topic models.
install.packages("newsmap")
install.packages("seededlda")
    Installing package into '/usr/local/lib/R/site-library'
     (as 'lib' is unspecified)
     also installing the dependencies 'RcppTOML', 'here', 'png', 'reticulate'
     Installing package into '/usr/local/lib/R/site-library'
     (as 'lib' is unspecified)
    Installing package into '/usr/local/lib/R/site-library'
     (as 'lib' is unspecified)
# Para dados preformatados, requerer os pacotes abaixo.
require(quanteda)
require(readtext)
# Setting WD (Working Directory) trhough code
# Checking actual WD
getwd()
# Setting properly the WD, if wanted by coding setwd() [replacing \ in a path with /]
# setwd("C:/Users/labo2/OneDrive/Documentos/ROOT BFreitas")
# Alternatively, you can use the readtext package to import character (comma- or tab-separated)
## values. readtext reads files containing text, along with any associated document-level variables.
#
library(readr)
> Loading required package: quanteda
     Package version: 3.2.3
     Unicode version: 10.0
     ICU version: 60.2
     Parallel computing: 4 of 4 threads used.
    See <a href="https://quanteda.io">https://quanteda.io</a> for tutorials and examples.
     Loading required package: readtext
    Warning message in library(package, lib.loc = lib.loc, character.only = TRUE, logical.return = TRUE, :
     "there is no package called 'readtext'"
    '/content'
#Carregando Coleta01
coleta01 <- read_csv("coleta01.csv")</pre>
View(coleta01)
```

```
Rows: 47282 Columns: 90

— Column specification

Delimiter: ","

chr (57): user_id, status_id, screen_name, text, source, reply_to_status_id...

dbl (18): display_text_width, favorite_count, retweet_count, quoted_favorit...

lgl (11): is_quote, is_retweet, quote_count, reply_count, symbols, ext_medi...

dttm (4): created_at, quoted_created_at, retweet_created_at, account_create...

i Use `spec()` to retrieve the full column specification for this data.

i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

user_id	status_id	created_at	screen_name	text	source	display_text_width	reply
<chr></chr>	<chr></chr>	<dttm></dttm>	<chr></chr>	Um homem negro de 60 anos foi espancado ATÉ	<chr></chr>	<db1></db1>	
x1250437852267073541	x1329768152125874176	2020-11-20 12:46:48	eimariainess	A MORTE por ASSASSINOS (seguranças) do Carrefour em Porto Alegre. Entenderam? UM HOMEM NEGRO DE 60 ANOS FOI ASSASSINADO POR CRIMINOSOS QUE TRABALHAM PARA O CARREFOUR!	Twitter for Android	140	
x1250437852267073541	x1329767543658123264	2020-11-20 12:44:22	eimariainess	Vídeo registrado no Carrefour Porto Alegre. O homem negro espancado, veio a falecer. Tudo isso no DIA DA CONSCIÊNCIA NEGRA. Esse é o Brasil que se comove com George Floyd e não percebe que essa merda acontece TODO DIA por aqui. https://t.co/zeEfmWU45E	Twitter for Android	140	
x739214299504234496	x1329768151685476358	2020-11-20 12:46:47	_juliamalagutti	divulguem apenas fotos dos assassinos que a carrefour contratou, mas uma vez essa empresa maldita https://t.co/MMM3Rf0y0l	Twitter for iPhone	139	
x1290069965597548545	x1329768150121013249	2020-11-20 12:46:47	Milene_Cassana	Não compartilhem o vídeo do assassinato brutal de uma pessoa negra dentro da Carrefour ontem, compartilhem o rosto dos assassinos.	Twitter for Android	140	
<pre>#library(readr) #Carregando Coleta01 #Criar pasta no driver no c coleta02 &lt;- read_csv("colet View(coleta02)</pre>		ou alterar	o caminho para um	caminho específico no	drive		
x1171921137774866433	x1329768149743448064	2020-11-20 12:46:47	xuecandyboy	funcionários, na semana em que falamos de consciência negra corpos negros são impostos a violência, ao desamparo, a luta antirracista não acabou. #VidasNegrasImportam #BlackLivesMatter https://t.co/yvRd3GRx1N	Twitter for iPhone	140	
x1074774467488960514	x1329768149730910209	2020-11-20 12:46:47	medellin201	FOGO NO CARREFOUR!!!	Twitter for Android	35	
x867217707179143168	x1329767754283540480	2020-11-20 12:45:13	aeugenios	Revoltante. No dia Nacional da Consciência Negra, no Carrefour de Porto Alegre um homem negro, é espancado até morte por segurança e PM. Enfrentar a cultura do racismo estrutural no Brasil é preciso resistência e luta.	Twitter for Android	140	

/06/25	i, 19:52		Cópia de c-	pia-de-atualizado_	qtda_bfreitas_r.ipynb - Colab			
<b>→</b>	Rows: 46426 Columns: 9		` for details		justificar o assassinato no Carrefour: "Se fosse um homem branco com Twit		440	
	<pre>dbl (20): display_tex lgl (11): is_quote, i</pre>	<pre>catus_id, screen_name, ct_width, favorite_coun .s_retweet, symbols, ex quoted_created_at, re</pre>	12:45:47 text, source, t, retweet_co t_media_type,	unt, quote_count, quoted_verified,	teria tido o mesmo Andro id desfecho?"  #VidasNegrasImportam  r #fogonosracistas te	for oid	140	
	<pre>i Use `spec()` to retr i Specify the column ty</pre>	ieve the full column spypes or set `show_col_t	pecification types = FALSE	for this data. ` to quiet this me	Um negro foi espancado até a morte em loja do ascarcefour, Porto Alegre. Supermercado e banco			
	user_id	status_id	created_at	screen_name	são catedrais da exclusão: um branco q <sup>text</sup> Twit	ter source	display_text	_widtl
	x867217707179143168	x13297681490179 <b>11296</b>	2020-11-20 1 <b>2.446</b> : <b>447</b>	aeugenios >		for <chr></chr>	140	<dbl:< td=""></dbl:<>
	x139555167	x1329833557108330497	2020-11-20 17:06:41	thiago_pmartins	Um homem negro espancado até a morte dentro do Carrefour e isso não abalou em absolutamente nada o desempenho da empresa no mercado financeiro. Não adianta dizer que vidas negras importam, enquanto o dinheiro que move o mundo segue ignorando a vida negra e tratando como lixo. https://t.co/xJwkjHtDy6	Twitter for Android		ΝŁ
		x1329768148531351553 x1329833557326491648	2020-11-20 2020-16-27 2020-16-27 17:06:41		povo i confeças o Usa da Consciencia Regra o Usa da Consciencia da Consciencia da Consciencia da Incompleta da Incompl	ter for oid Twitter for Android	259	Νŧ
head	ualizao cabeçalho dos d (coleta01)	x1329833557439770625 latasets	2020-11-20 17:06:41	c_wolff02	→funcionário faleceu e esconderam o corpo com guarda-sol para manter o mercado em funcionamento.  →agressão a homem negro e com deficiência física →cachorro morto com extrema brutalidade →homem negro espancado e morto Esse é	Twitter Web App		ΝŁ
head	(coleta02)				racial. Em homem negro foi			
	x358905632 x227393613	x1329768148090937344 x1329833557515247622	2020-11-20 2049-16-39 17:06:41	ToolooFaloo	especial aliminates por presulting shaking use no confluint perfect of the state our. Of the burner of the state of the st	ter eb Twitter for PP Android	140	N/
	x1157683243531022337	x1329833558198906882	2020-11-20 17:06:42	freudexplica	A CADA 23 MINUTOS UM JOVEM NEGRO É ASSASSINADO NO BRASIL. A CADA 23 MINUTOS UM JOVEM NEGRO É ASSASSINADO NO BRASIL. A CADA 23 MINUTOS UM JOVEM NEGRO É ASSASSINADO NO BRASIL. A CADA 23 MINUTOS UM JOVEM NEGRO É ASSASSINADO NO BRASIL. O DE MINUTOS UM JOVEM NEGRO É ASSASSINADO NO BRASIL.	Twitter for iPhone		Νź
	x1237803190583799811	x1329833558249181185	2020-11-20	Rosa28123955	no dia da consciencia negra um homem é 😜	Twitter for		N/
no:lloc	olah research google con		17:06:42		espancado e morto	Android		4/13

	_
•	
7	•

<del>→</del>					llm homom norre loão			
J	user_id	status_id	created_at	screen_name	Um homem negro, João Alberto Silveira <b>taxt</b> tas <b>séerce</b> 40 anos morreu espancado	display_tex	t_width r	eply_to_:
	<chr></chr>	<chr></chr>	<pre><dttm> 2020-11-20</dttm></pre>	<chr></chr>	por seg@hanças d6hr>	Twitter for	<dbl></dbl>	
	x1250437852267073541	x1329768152125874176	2020-11-20 12:46:48	eimariainess	Um homem negro de 60 anos foi espancado ATÉ  A MORTE por ASSASSINOS (seguranças) do Carrefour em Porto Alegre. Entenderam? UM HOMEM NEGRO DE 60 ANOS FOI ASSASSINADO POR CRIMINOSOS QUE TRABALHAM PARA O CARREFOUR!		140	
	x862338664021577729	x1329768147277258753	12:46:46	TalinhosPlays	falannguenne reges socials control of billion of social of the control of the con	or ne	228	x132976
	x1250437852267073541	x1329767543658123264	12:44:22 2020-11-20	eimariainess	NEGRAA Essemble Blastlomem for que se cappaye posta la parcio	Twitter Web	140	
	x1323651498585317381	x1329833558853218304	17:06:42		percebgiguecessenmendespeito acontemps@D&AWAPOv0f7cb	Twitter Web App		N/
	x1038967723005489152	x1329768147189178371	2020-11-20	melizaantunes	https://l.cs/acethal/article/ades Twitt W divulguem apenas fotos dos assassinos que a	er eb +:	98	
	x739214299504234496	x1329768151685476358	2020-11-20 12:46:47	_juliamalagutti	carrefour contratou, mas uma vez essa empresa maldita https://t.co/MMM3Rf0y0l		139	
		x1329768146929131521 x1329768150121013249	2020-11-20 12:46:46 2020-11-20	HenrikShirley	em que raiamos de consciente de la consc	er eb op	140 140	
		x1329833559419482116	20 <b>20:46:20</b> 17:06:42	Milene_Cassana beloti_is	a Carrendarion Englimentrold  Carrendarion Engliment Carrendarion  Carrendario Engliment Carrendarion  Carrendario Engliment Carrendario Engliment Carrendario  Carrendario Engliment Carrendario Englimen	Twitter for Android	140	NA
	x778709674124128256	x1329768150032916480	2020-11-20 12:46:47	TrincaGuilherme	Tem que arrebentar todo o carrefour Android  o assassimato brutar นธรระ  Umhbศาลศารศาสมาร์ ให้ก duro gสตุลจะมาสงระธากละลo.		35	
	x881460697	x1329833559406878721	2020-11-20 17:06:42 2020-11-20	dinopon	Qu <b>e atterézat É</b> opreciso funci <b>r</b> ahban roassos cas que tii # <b>vida</b> kalograsan pertam . conégiêneisana rambiayitter	Twitter for Android		N/
	binando as duas coleta a0102 <- rbind(coleta0							
					апшта <b>рпярь:</b> !/чя:/есячиничкLULZ			
corpu corpu names	s_BFreitas <- corpus(c s_BFreitas (corpus_BFreitas)	· -	esmo (criamo	s de cada um caso	o haja a necessidade de algum	a analise te	mporal)	
head(	corpus_BFreitas)	x13299335585158 <b>6976</b>		Hendikshayev	estabelecilijenius k seus (ah wili b functionarios vao assimicada W.	ei i witter vveb	<dbl></dbl>	N/
	x93258006 <b>2704144384</b>	x132976758027868160Y	17:06:43	Henrik Shirlev	Um homem negro espancado até a morte	eb ∆nn	140	147

x932580062704744384	x1329983588578889689	17:06:42	Henrik Shirlev funcionarios vão assumir de Web App 140	N/
x139555167	x1329833557108330497	2020-11-20 17:06:41	Um homem negro espancado até a morte dentro do Carrefour e isso não abalou em absolutamente nada o desempenho da empresa no mercado thiago_pmartins thiago_pmartins thiago_pmartins  Twitter financeiro. Não adianta dizer que vidas negras importam, enquanto o dinheiro que move o mundo segue ignorando a vida negra e tratando como lixo. https://t.co/xJwkjHtDy6	
		12:44:22	'porra a giopo ta que ta Configuação Diagneta do há Configuação Diagneta do di	
x761734926215606272	x1329833560220594176	2020-11-20 17:06:42	andreiapmirum preto assantinem e Twitter for assantinem regression e Twitter for espancado atessassinado. Android	N/

```
no caparocalmos tada a
no supermercado seguenwitter
→ text1:
             xt1:
x793Um homen degro de 60 angs foi espancado ATÉ 20000时程中or ASSASSINOS (seguran中的明新语句 电影图 dofor do Carrefour em Porto Alegre. Entenderam? UM HOMEN6MEGRO DE 60 ANOS FOI algunadia rental allegre. Entenderam? UM HOMEN6MEGRO DE 60 ANOS FOI algunadia rental allegre. Entenderam? UM HOMEN6MEGRO DE 60 ANOS FOI algunadia rental allegre.
                     ASSASSINADO POR CRIMINOSOS QUE TRABALHAM PARA O CARREFOUR!
                                                                                                                                                                        gente precisa é de
                                                                                                                                                                    considiênician le mangro de 60
                     "Vídeo registrado no Carrefour Porto Alegre. O homem negro espancado, veio a fale@ntancespaguado ATÉ A
                     isso no DÍA DA CONSCIÊNCIA NEGRA. Esse é o Brasil que se comove com George Fander Espesia SASSINOS
                     não percebe que essa merda acontece TODO DIA por aqui. https://t.co/zeEfmWU45E' (segutançan)áticiSarrefour
                                                                                                                                                                #VidasNegras lemp protono Alegre.
          text3:
                                                                                                                                                                                                                                  Twitter for
                     2020-11-20 **Trivûlgûêmî apênasîfotos dos assassinos que a carrefour contratou, más uma vez essa
                                                                                                                                                                    ÷funcionário faleceu e
                     empresa maldita https://t.co/MMM3Rf0y0l
                                                                                                                                                                     esconderam o corpo
          text4
                    'Não compartilhem o vídeo do assassinato brutal de uma pessoa negra dentro da Carrendiguarda-sol para
                                                                                                                                                                   manter o mercado em
                     ontem, compartilhem o rosto dos assassinos.'
                                                                                                                                                                             funcionamento.
           text5
                                'Tem que arrebentar todo o carrefour' text6:
                     'Um homem preto foi assassinado no Carrefour por funcionários, na semana em que Talarressão a homem
                     de consciência negra corpos negros são impostos a violência, ao desamparo, a luta antirracista pegro e com
                                                                                                                                                                                                             Twitter
              x97660363636347246045N9gresignostergy#BlackLive9Ratter4ftps://t.co/yvRd3Gfbx1N'
                                                                                                                                                                            deficiência física
                                                                                                                                                                                                                                                            NA
                                                                                                                                                                    →cachorro morto com
                                'FOGO NO CARREFOUR!!!' text8:
                     'Revoltante. No dia Nacional da Consciência Negra, no Carrefour de Porto Alegre um extrema brutalidade
                                                                                                                                                                              →homem negro
                     negro, é espancado até morte por segurança e PM.\n\nEnfrentar a cultura do racismo
                                                                                                                                                                       espancado e morto
                     estrutural no Brasil é preciso resistência e luta.'
                                                                                                                                                                     Esse é o histórico do
          text9:
                    'Aos FDPs que tentam justificar o assassinato no Carrefour:\n"Se fosse um homem brander GUANDO? © QUANDO?
                     \'cara de gente de bem\', teria tido o mesmo
                                                                                                                                                                https://t.co/EQlgCp81Kj
                     desfecho?"\n#VidasNegrasImportam\n#fogonosracistas'
                     Um homem negro foi
'Um negro foi espancado até a morte em loja do Carrefour, Porto Alegre. Supermercado até a morte em loja do Carrefour, Porto Alegre. Supermercado até a morte em loja do Carrefour, Porto Alegre. Supermercado a de consciención de la conscienci
             139
                    no supermercado, 18hs:

1:

'Não compartilhem o vídeo do assassinato brutal de uma pessoa negra dentro da Carretour negro foi especialde de la compartilhem o rosto dos assassinos de 2020-11-20 speakinow #Vides bassassinos de 11-20 speakinow #Vides bassassinos 
          text11:
                    x111月1689943891359597 0.19598495857515247622
text12: x227393613 x1329833557515247622
                                                                                                                                                                                                                                   Twitter for
                   NA
                                                                                                                                                                                                                                      Android
                                                                                                                                                                                                                                                            NA
          text13:
                     diversas vezes. https://t.co/OSJZXsXVrU'
                                                                                                                                                                           vídeo daqueles noientos
                                 'Deus que me perdoe mas , FOGO NO CARREFOUR https://t.co//Evm1M2Mkvctext45:3 MINUTOS
          text14:
                     'Repudiamos qualquer ato de discriminação racial. Em pleno 2020, o racismo ainda está JOVEM NEGRO É
                     presente e precisa ser combatido com urgência. O futebol tem o dever de participar desas lutas INADO NO
                     por um mundo mais respeitoso, solidário e igualitário. #vidasnegrasimportam
                                                                                                                                                                     BRASIL. A CADA 23
                     #diadaconsciêncianegra https://t.co/3SkCVUE8zf
                                                                                                                                                                 MINUTOS UM JOVEM
          text16:
                     'Carrefour querendo a qualquer custo superar a Havan como marca mais odiada de quangañas INADO NO
                                                                                                                                                                                                             Twitter
                     segue a bojada.
                                                                                                        2020-11-20
                                                                                                                                                                     BRASIL. A CADA 23
           text 67683243531022337 x1329833558198906882
                                                                                                                                   freudexplica_
                                                                                                                                                                                                                                                            NA
                                                                                                                                                                                                                    for
                     'O Carrefour no dia da Consciência Negra. \n\n#carrefourRacista #CarrefourAssassino
                                                                                                                                                                 MINUTOS UM JOVEM
                                                                                                                                                                                      NEGRO É
                     #VidasPretasImportam https://t.co/8p2VK56PWj'
                                                                                                                                                                     ASSASSINADO NO
          text18:
                                                                                                                                                                     BRASIL. A CADA 23
                     'no dia da consciência negra um homem é espancado e morto dentro de um supermentado DI UM JOVEM
                     Carrefour. existe mesmo consciência? todo dia um Beto morre e a gente nem sabe porque
# Criando uma matriz a partir do corpus, a partir do qual vamos conseguir fazer as análises
corpusBF.dfm<-dfm(corpus_BFreitas,</pre>
                                  remove_punct = TRUE,
                                  case_insensitive=TRUE,
                                  remove = stopwords("portuguese"),
                                  verbose = TRUE)
colnames(corpusBF.dfm)
corpusBF.dfm
head(corpusBF.dfm)
              AUDULUUUTUUU LUUUUUU A LUEUUUUUUEU LTULTUUU
                                                                                                             17:06:43
                                                                                                                                                                                                                                      Android
                                                                                                                                                                           nada#edizeiene#Negra
                                                                                                                                                                       eanted the pagarrafals um
                                                                                                                                                                       aés de grorenda mos dumos
                                                                                                                                                                      minutedportingspoiedade
                                                                                                                                                                            as pessoas achamismo.
                                                                                                                                                                          padamatanegrataripantam Twitter
                                                                                                        2020-11-20
                               x4268899343 x1329751804926636033
                                                                                                                                                   hlastikk
                                                                                                                                                                     nela gente e invadir total
                                                                                                                                                                                                                                                                  274
                                                                                                                                                                       Acordei no ódio hoje vendo
                                                                                                         2020-11-20
                                                                                                                                                                                                                                   Twitter for
                                 x201556631 x1329833562523246593
                                                                                                                                              Fehsimoees
                                                                                                                                                                                        esse cara sendo
                                                                                                             17:06:43
                                                                                                                                                                                                                                       iPhone
                                                                                                                                                                            espancado no carrefour
                                                                                                                                                                            וססט אין ט טומווטט אמו
                                                                                                                                                                    continuance ศายการอุดเนอาdo
                                                                                                                                                                       aconteceuté aforobrajamais
                                                                                                                                                                       nos calariamos com o caso
                                                                                                                                                                       ægouiÇamtefourrjáhlamojem de
                                                                                                                                                                          unco anotas fadialondotadomente
                                                                                                                                                                    lamenta o aossoarsistion aed opupeor 2
                                                                                                         2020-11-20
                                                                                                                                                                                                                                   Twitter for
                                   x72432728 x1329833563416629251
                                                                                                                                                  luiz RDL
                                                                                                                                                                     vaietourances vidécosias four... Twitter
                                                                                                                                                                                                                                                                                       NA
                                                                                                        2020:06:20
                                                                                                                                                                                                                                      Android
                                 x304050626 x1329751800531005440
                                                                                                                                                                            tanto dhosoasropopubejénais
                                                                                                                                            Neverton77
                                                                                                              11:41:49
                                                                                                                                                                             anccontencer.atião mãodenAndroid
                                                                                                                                                                         todetxamogravistênovencer
                                                                                                                                                                     atélbonjie a hão resposica Adoberto
```

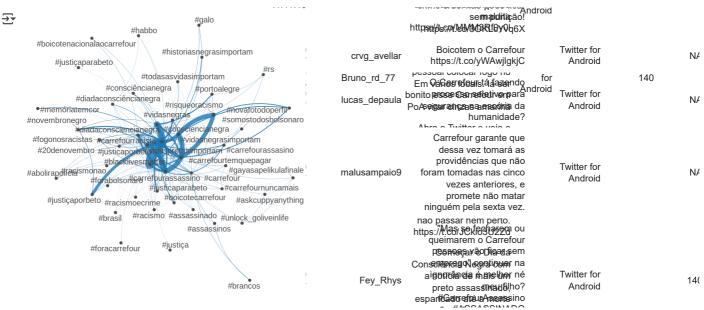
	•		. –				
₹	Creating a dfm from a corpus input			#Vidas <b>Neguæslmajs?t</b> am https://t.co/hSjkq6EOni o Brasil não tem um			
	Warning message: "Xdf61755685594472960prxt329633563752664282)	, 2020-11 <del>,</del> 20 first 17:06:43	PedroCamposs2	cancelem o Carrefour	Twitter for iPhone		N/
	Creating a tokens object from a corpus input.	• • •		reuro, Luana barbosa, A <b>ூithię கேர்</b> ஜ்i <b>Evaldja</b> eleTwit	tter .		
	Warning message: "case_insensitive argument is not used." starting tokenization	2020-11-20 11:41:49	jaquebelieber6	queatovióผุกคาญisa Carolinal O ปฏิทธิโด้ Minge Les istandr NOMES PALINDANque	for	140	
	text1 to text10000 x1206667052888530945 x1329833563995430916 preserving hyphens	2020-11-20 17:06:43	Astrio78855248	ছাহাৰত justificers in rotto de #Vidas Negristino bradeixa পুমারটের মান্টের চিন্ত কি ber que infelizmente não é um	Twitter for Android		N/
	$\ldots$ preserving social media tags (#, @)			e <b>sas</b> céi <b>soflaidoc∤e</b> sque ta <b>ndsássinfælizhaeloté</b> onão			
	segmenting into words			Alberto Silveir <b>sersetati</b> tatimo dentro do CARRISTICAPORSeto			
	text10001 to text20000			Carrefour já vem	Hor		
	preserving hyphens			assassinando vidas há um tempo. Parem de divulgar			
	preserving social media tags (#, @)			fotos do Beto morto, vídeos dele sendo			
	x10 <b>89g490†\$8g6662274</b> 0r\$\$329833563995430913	2020-11-20 17:06:43	inejweasley	espancado e humilhado, comecem a mostrar a cara	Twitter for Android		N/
	text20001 to text30000	17.00.43		dos assassinos e não da vítima em estado de morte,	Allalola		
	preserving hyphens			no chão, sangrando. #justicaporbeto			
	preserving social media tags (#, @)			#VidasNegrasImportam https://t.co/fFjOYErTFk			
	segmenting 1176 Work 1329751799985819651	2020-11-20 11:41:49	FilipeCifali	DEUSEUL HIUND DEUL EUL	/eb	280	
		2020-11-20		pensealmante intratem,  vantas interes dizersque	npp Twitter Web		
	x1289053241632884224 <sup>40</sup> 29329833564234461184 preserving hyphens	17:06:43	kinhakinn	de <b>redicteria. Videar iengari</b> am. forma de mu <b>tilalatia</b> , a <b>Beja</b> leto	App		N/
	preserving social media tags (#, @)	:	:	ag <b>lattes∷∜bea/in∃a</b> Ce6ejwZ :	:		
	segmenting into words	•		a morte de João Alberto	•		
	text40001 to text50000			Silveira Prettas, Satomendo NEGRO, ASSEASSINAPOS de gustos Banhardos em			
	·x797059525075003136° x1329845864286867458	2020-11-20 17:55:36	NaoTaSenduFacil	ASITIA (AD LOT PERSENSE NO Por Segurially a Euleria	Twitter Web		N/
	preservi283592182 ms4329758799578980352	2020-11-20 11:41:49	eder_organica	CARREFOUR de Porto en la companya de compa	for	140	
	segmenting into words			dia da Consella de Pho NEGRA, é revolta fite aro.	nie		
	text50001 to text60000			A cidade de Porto Alegre			
	preserving hyphens	2020-11-20		está abandonada. Falta políciamento para a	T 10 1		
	x107246996 x1329845864626610183preserving social media tags (#, @)	17:55:36	kassimentz	cidade, mas para a segurança de um supermercado tem.	TweetDeck		N/
	segmenting into words			https://t.co/yavPCirl9r			
libr	<pre>lvando o metadado em uma tabela ary(MASS) ("corpusBF.dfm", file = "corpusBF.dfm")</pre>						
	***************************************	2020-11-20		#20de/ <b>ମ</b> ଧ୍ୟକ୍ରମାଧାର =			
-	ais são as hashtags? dfm.corpusBF <- dfm select(corpusBF.dfm, patte	ern = ("#*")	1				
cu <sub>B</sub> .	arm-serves varm_serve(eorpassrrarm, parce	i / .ບບ.ວບ	,	https://1.26/14/2592A37W4u=	Whh		
#htt #Lis topt	rega as n hashtags que mais ocorrem no texto ps://quanteda.io/reference/topfeatures.html t the most (or least) frequently occurring fea ag.corpusBF <- names(topfeatures(tag.dfm.corpu(toptag.corpusBF, 50)		dfm, either as a w	hole or separated by docume	ent.		
<b>→</b>	"#vidasnegrasimportam" · "#carrefourassassino" · "#vidas "#diadaconsciencianegra" · "#blacklivesmatter" · "#consci "#justiçaporbeto" · #carrefourracista' · "#g1' · "#boicoteca "#vidasnegras' · "#novembronegro" · "#portoalegre' · "#fourafotodopeffil" · #diadaconsciêncianegra' · "#joaquil "#historiasne (1888) / 1888	iencianegra' · '# arrefour' · '#justi gonosracistas'	#carrefour' · ˈ#risqueoɪ icaparabeto' · ˈ#brasil' · ˈ#carrefourassasino'	(seguranças) do Carrefour	Twitter for iPhone		N/
save	fcm.corpusBF <- fcm(tag.dfm.corpusBF) ("tag_fcm.corpusBF", file = "tag_fcm.corpusBF" ames(tag_fcm.corpusBF)	')					
	X161901532 X1329751798190645248	11:41:48	AliomarCamara	novamente o Carretour 'Não rexiste lacismo no Andr	tor oid	135	
	x1082670152255266816 x1329845865457082371	2020-11-20 17:55:36	Diowsoares07	episodiosi de iz de da após morte de la para no Carrefour	Twitter for Android		N.
				O @https://diable.com/DsiyfilaBBq5			

0,00,20	, 10.02		oopia ao o p	ola ao ataalizaao_	qua_bironao_i.ipyrib oo	ias		
<del></del>	"#vidasnegrasimportam" · "#bla" "#carrefourracista" · "#carrefou" "justicaporbeto' · "#racismo' · "#consciencianegra' · "#justicaporpaoalbeto' · "#racismo' · "#chegadepreconceno · "#racismo' · "#carrefourassasino' · "#carref '#forabolsonaroeseubandodee' redhsenado' · "#portoalegre' "#oportunidades' · "#vamosjun	rassassino' · #vidaspretasir '#carrefour' · #vidasnegras' porbeto' · #justicaparabeto' ismonāo' · #g1' · #extraass 1304841565608 · #solioteca ournao' · #morganfreeman' criminososecorruptos' · '#for · '#rs' · '#boicoteaocarrefour'	nportam' · '#di · '#novafotod · '#rádiogaúcl   3230401 - 24000  rrefio 65:346coi · '#blm' · '#for  acarrefou' · '# - '#diversidad	iadaconsciencianegra operfil' · #risqueorac ha' · #carrefourtemque deriasereu' · strino nsciêncianegori acarrefour' · sosamapa' · #brasil' de' · #igualdade' ·	a' consciência negra. isninomento de comemorar o uepagas conquistamos. Mas, também, dia de dizer que não vamos parar. Vou morrer lutando.	Twitter Web App		N/
	"#eliminacaoafazenda' · '#just '#assassinato' · '#assassinato '#notoracism' · '#brasilantiraci '#pcdob65' · '#comunistarj' · '#atg265807.1\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	carrefour' '#boicote' '#mul smo' '#acabecomaculturara '#comunistas' '#igualdadea '#A39A458556249544920der obioto' '#mucb' '#vidane '43297531624936169489rrefo carrefour' '#blackblocks' '# imostodoatoracista' '#lolerš	rder' · '#cowar acista' · '#negr acial' - 1 · #5ahia iovengg 6 · #1 ggaappd 2mpa uur 1 · #sapwat #habbo' · #tbt uncia' · '#impa	rd' · '#boycottcarrefouros' · '#pcdobrj' · '#pc anoticias' · '#scc10' · ' fenfakutjawa#eisasi z' · '#feriado' ch' · '#celtsal-miejiti · '#racismonao' · tia' · '#betoeterno' · '#	daniegra começa AB BRASEste #negraparimentado y alamais nonembro y del de la mais nonembro y del del mais madares y del del del mais m	101	140	N/
	"#racista' · "#noracism' · "#che "#conciencianegra' · "#podcas "#podcastsbrasileiros' · "#orque "#racistyleiros' · "#totalez "#feminicidios' · "#totamericano' "#assholes' · "#lutoamericano' "#raça' · "#direitos' · "#racismo	tbrasil' · '#podcastpt' · '#pod lholgbt' · '#lgbtia' · '#psol' · ' <u>½</u> 13 <sup>2</sup>	castbr' · '#lgbt <b>/gogtupezg</b> bra af <b>azegsla</b> &2' · oaoracismo' · no' · '#negra' ·	podcasters' · '#podos ısil' · '#politica' · '#coทริศิทิปิยิ <b>ฮิยิเลลิ</b> r '#brazil' · '#terezadet '#memória' · '#resisto	sfera' . consciencia negra, um homem preto morreu <sub>rid'</sub> . espancado por penត្សិន្ត្បម្ភាឌពាças" do carrefour	Twitter for Android		N/
	"#carrefourcancelado' · '#kkka "#assassinos' · '#naocompren "#paíszinhodebosta' · '#bolsor "#paremdenosmatar' · '#revolt "#greys' · '#violencia' · '#brasil "#quadrilha' <u>v</u> ##953986645 '#k "#genocidionegro' · '#racismon "#acmarca' · '#racismonecrimel as mais frequentes? at_fcm.corpusBF <- fcm_se	urrefour' · '#ad' · '#homicídioc ocarrefour' · '#todesmerecer narismodoentio' · '#revolta' · ante' · #verdadedosfatos' · ' racista' · '#ridaymorning' · '# into299469674359932820' · ' mata' · '#atéquando' · '#bolso pediondo' · '#michallenhama	loloso "#crue mrespeito" '# #pcdobbahia #sextou' '#re lifeggesonwit #bjesib 36 onaroracista'	eldade' · '#repúblicad carrefourclose' · 65' · '#ifpi' · '#extra' · espeito' · '#greysanat hbts' · '#unlock_goliv abols <b>masdes@pad</b> a ' '#premiosinnovaciór a ds' · '#a and all relation	compartition racismo, nem compartition racismo, nem what is sent to the compart of the compart o			N/
	("topgat_fcm.corpusBF", f		-	cop cag veo. pass. 7				
head	<pre>(topgat_fcm.corpusBF)</pre>							
<del>_</del>	XZO173040 X Feature co-occurrence ma	ເວລະອວ4ວວວາວບາດບາດວວວວດ ntrix of: 6 by 50 features	ne\$7:55:36	піетіеуетіго	e ieva irivertida de Constantino	iPhone		INF
	features	#vidasnegrasimportam =	#blacklives	matter	https://t.co/YEw6U5FAmH			
	x#yjdasnagfaðimportam x#blacklivesmatter	1329845867931725826 <sup>14</sup> 2	2020-11-20 17:55:36	1658 Tamiih89797721	#VidasNegrasImportam #VidasNegrasImportam Homem NEGRO é	Twitter for Android		N/
	<pre>#fogonosracistas #diadaconsciêncianegra</pre>	0 1		0 0	Bom dia Esquerda 💜 🟲			
	#carrefourracista	0		0	"Eu fico com a pureza da			
	#carrefourassassino	0		0	resposta das crianças É			
	_	features		_	a vida, é bonita e é bonita" O racismo é coisa			
	features #widesneggesimpontem	#fogonosracistas #diad 75	daconsciênc:	ianegra 82	de gente grande e			
	#vidasnegrasimportam #blackli <b>%29005</b> 24526 x		2020-11-20	1.5	estúpida, incompetente e	Twitter for		N/
	#fogonosracistas	1329043000093227701	17:55:36	frangaieta	idiota. SOMOS TODOS	Android		INF
	#diadaconsciêncianegra			0	IGUAIS #VidasNegrasImportam			
	#carrefourracista	0 0		0	Chega de naturalizar a			
	#carrefourassassino	features		0	barbárie contra nossos			
	features	#carrefourracista #ca	rrefourassa		irmãos! 🥮 🏲			
	#vidasnegrasimportam	220		992	https://t.co/TA11YL1L1B			
	<pre>#blacklivesmatter #fogonosracistas</pre>	46 4		410 68	anga foi espande do ATÉ do			
	#diadaconsciêncianegra			19	assassinato do nomem neରୁ ଚନ୍ଦିର ଅଧାନନ୍ତିour.			
	#carrefourracista	0		320	Espain Sall Faring Solds. os			
	#carrefourassassino	features	2020-11-20	12	ROSTE OUTSERS ASSENCE.		4.40	
		ተ <b>ያደ፬ ኮሪ የ ፣</b> 97255315457 ለተ <b>ያደ ያቆ4 5 8 6 8 8 5 3 1 ሰ ያ 0 6 2</b> am 3 562	02 <b>0-11-20</b> # <b>d18dac01</b> sc: 17:55:37	_jenifer0 iencian <b>agad</b> antasa 1164	Alegre. Fatenderam holle de HOMEN DE GRODE GAN har	for Twitter for ndroid Android	140	N/
	#blacklivesmatter	201		167	engajamehti 9 Nã5 V um cassassas Salvas Podio Pacial			
	#fogonosracistas	7		53	CRIMINGSOS QUEque			
	<pre>#diadaconsciêncianegra #carrefourracista</pre>	14 117		3 72	TRASSOL HAM PREBAStem			
	#carrefourassassino	875		265	que fazer e ouvir e apoiar.			
	Continue	features			Toda solidariedade e			
	<pre>features   #vidasnegrasimportam</pre>	#justicaporbeto #raci: 710	smo 22		respeito à família e amigos			
			20220-11-20	huaisangfan	de João Alberto Silveira Freitas e à toda a	Twitter for		N/
	#fogonosracistas	62	<b>1</b> <sub>0</sub> 7:55:37	naaisangian	comunidade negra.	Android		1 1/
	<pre>#diadaconsciêncianegra #carrefourracista</pre>		0		#vidasnegrasimportam			
	#carrefourracista #carrefourassassino	51 521	6 10		#historiasnegrasimportam			
	[ reached max_nfeat				(que já Aprosums lais somi a no			
	x756146786712989697 x		2020-11-20	PortalVarada	de Casore feoinalen Rostrou Anlegre,	WordPress.com		N/
	710014010011280808/ X	1023040010000333000	17:55:37	i-ortarvarada	Mourão nega racismo	vvoiuriess.com		INF
					Um homen negro de 60 anos foi espancado ATÉ			

- # Tive problemas com a função textplot\_network(). Aparece erro.
- # Por isso, abrir https://cran.r-project.org/web/packages/quanteda.textplots/quanteda.textplots.pdf para informações completas sobre a dibrary("quanteda.textplots")
- # Depois de pesquisar, li que a função textplot\_network tinha sido movido para o pacote quanteda.textplots, conforme p. 78, Package 'quatextplot\_network(topgat\_fcm.corpusBF,

min\_freq = 0.1,
edge\_alpha = 0.8,
edge\_size = 4)

# Para mais informações sobre network, ver <a href="https://www.kateto.net/wp-content/uploads/2015/06/Polnet%202015%20Network%20Viz%20Tutorial%">https://quanteda.io/articles/pkgdown/examples/twitter.html?q=network</a>



# Nuvem de hashtags:

set.seed(132) #random

textplot\_wordcloud(tag.dfm.corpusBF, min\_size = 0.5, max\_size = 4, min\_count = 3, max\_words = 150, color = "darkblue",
 font = NULL, adjust = 0, rotation = 0.1, random\_order = FALSE, random\_color = FALSE, ordered\_color = FALSE,
 labelcolor = "gray20", labelsize = 1.5, labeloffset = 0, fixed\_aspect = TRUE, comparison = FALSE
)

# Apareceram avisos sobre tamanho e outras configurações do resultado gerado. O comando abaixo disponibiliza os 50 primeiros. Depois vej warnings()

# Diminuí de 500 para 150 o 'max\_words'. Os warnings diminuíram.

Warning message in wordcloud(x, min\_size, max\_size, min\_count, max\_words, color,um cachorro com uma "#vidasnegrasimportam could not be fit on page. It will not be plotted." barra de ferro -agosto Warning message: desse ano que um In normalizePath("~") : path[1]="/sbuild-nonexistent": No such file or directory funcionário morreu e esconderam o corpo com+ calados, vamos expor os Twitter for FOGO RESERVE POWE!!! drymorais66 NA Android #VidasNegrasImportam #VidasNegrasImportam Twitter for HerosMoraes #CarrefourAssassino NA iPhone https://t.co/Mg7MWWG7DZ com a morte de um
O Carrefour tratando o
homem preto espancado
acontecimento como casoTwitter
por segurancas do
isolado so permina a for
negliciancia e hipocrisiAndroid
represents brazil more
than soccer and samba
than soccer and samba
#valinais da empresa Só
#vidasNegrasImportam,
hitidaticamente dois casos 140 nixcolaee #vidaspretasimportam carrefourassassino# Twitter for hvenizus NA Android ATENÇÃO: movimento stas. convocam protesto para chuvangarananapdechar o Carre estabelecimento Twitter lucasrcdriques 140 https://t.co/\DGUYDuig7

L\_Fotossintese

Tu és minha mais antiga

r ia maiar am nomom

paixão, vasco

# Justicanara Rota

Twitter for

Android

# Vamos analisar os tweets. Primeiro, devemos transformá-os em texto:

v\_0102 <- coleta0102 #parei aqui, 21/08/2022.

x1183615780040118272 x1329845874088943619

- # Neste ponto, estou buscando as correlações entre as palavras rede de palavras.
- # Encontrei um artigo: <file:///C:/Users/labo2/OneDrive/Documentos/R/1859-Texto%20do%20artigo-3983-2-10-20210608.pdf>.

ZUZU-11-ZU

17:55:38

- # Manejando o corpus0102 (as duas coletas de dados ao mesmo tempo)
- v\_0102.lower <- char\_tolower(corpus\_BFreitas)</pre>

ΝÆ

```
#https://tutorials.quanteda.io/basic-operations/tokens/tokens/
#https://quanteda.io/articles/pkgdown/examples/chinese.html
stop_words <- stopwords("portuguese")</pre>
v_0102.tokens <- tokens(v_0102.lower, remove_punct = TRUE) %>%
                   tokens_remove(pattern = stop_words)
v 0102.word <- v 0102.tokens %>% as.character()
                                                      2020-11-20
                                                                                         empresas que utilizam os
                                                                                                                       I WITTER TOP
       V7EE00E4470200E4042 V422004E074202E40004
                                                                                                                                                   NI/
#Bag of Words
bag_of_words \leftarrow dfm(v_0102.word, remove_punct = TRUE, remove = stop_words)
                                                                                         O que eu vi no Carrefour
→ Warning message:
     "'dfm.character()' is deprecated. Use 'tokens()' first."
                                                                                               foi um ser humano
     Warning message:
                                                                                            agredindo até a morte
     ": 1545418011d_n8t_be_used_for_tokens() argumen2020.44e20tokens() first." warning message: 17:55:38
                                                                                               outro ser humano.
                                                                                                                       Twitter for
                                                                                        Falhamos como espécie e
                                                                                                                         Android
     "'remove' is deprecated; use dfm_remove() instead"
                                                                                       ninguém fala sobre isso 💔
                                                                                        Todas as vidas importam...
View(bag_of_words)
Document-feature matrix of: 1,648,129 documents, 31,860 features (>99.99% sparse) Porto Alegre, 1971: o
     and 0 docvars.
                                                                                       o 20 de novembro para
             features
              homem negro 60 anos espancado morte assassinos seguranças carrefour celebrar o dia nacional
     docs
                                                                                         da consciência negra.
                        0 0
                                            0
                                                   0
                                                               0
       text1
                                0
                                                                           0
                                                                                      Porto Alegre, 2020: João
                        1 0
                                            0
       text2
                  0
                                 0
                                                   0
                                                       2020-11-20
11:37:54
                                                                           0
                                                                                                               Twitter
                                                                                         Alberto é brutalmente
       text3
                x4@15551@881 x132975081645@81805@
                                                                           0 kkkeviin
                                                                                     0
                                                                                                                                       140
                                                                                             assassinado por
       text4
                  0
                        0 0
                                            0
                                                   0
                                                                           0
                                                                                      0
                                                                                                               iPhone
                                 1
                                                                                               seguranças do
       text5
                  a
                        0 0
                                 0
                                            1
                                                   0
                                                               0
                                                                           0
                                                                                      0
                                                                                          @carrefourbrasil. No
       text6
                  a
                         0
                            a
                                 0
                                            a
                                                               a
                                                                           0
                                                   1
                                                                                      mesmo 20 de novembro.
            features
                                                                                       Nossa tragédia não tem
     docs
              porto
       text1
                                                                                      #VidasNegrasImportam?
       text2
                  0
       text3
                  0
       text4
                  0
       text5
                  0
       text6
     [ reached max_ndoc ... 1,648,123 more documents, reached max_nfeat ... 31,850 more
     features 1
```

## Duende goes on

```
View(v_0102.tokens)
Tokens consisting of 93,708 documents.
      text1 :
      [1] "homem"
                                          "60"
                           "negro"
                                                                        "espancado"
     [6] "morte"
[11] "alegre"
                           "assassinos" "seguranças" "carrefour"
                          "entenderam"
      [ ... and 8 more ]
     text2 :

[1] "vídeo"

[6] "homem"

[11] "tudo"
                           "registrado"
                                          "carrefour"
                                                         "porto'
                                                                         "alegre'
                                          "espancado"
                                                                         "falecer"
                           "negro'
                                                         "veio"
                           "dia'
      [ ... and 14 more ]
      text3:
       [1] "divulguem"
                                          "apenas"
       [3] "fotos"
                                          "assassinos"
       [5] "carrefour"
[7] "vez"
                                          "contratou'
                                          "empresa"
       [9] "maldita"
                                          "https://t.co/mmm3rf0y01"
      text4:
       [1] "compartilhem" "vídeo"
[6] "negra" "dentro
                                               "assassinato"
                                                               "brutal"
                                                                                  "pessoa"
                             "dentro"
                                               "carrefour"
                                                               "ontem"
                                                                                 "compartilhem"
      [11] "rosto"
                             "assassinos"
      text5:
     [1] "arrebentar" "todo"
      text6:
       [1] "homem"
[6] "semana"
                             "preto"
                                               "assassinado"
                                                                "carrefour"
                                                                                 "funcionários"
                             "falamos"
                                               "consciência"
                                                               "negra"
                                                                                 "corpos"
      [11] "negros"
                             "impostos"
      [ ... and 8 more ]
      [ reached max_ndoc ... 93,702 more documents ]
```

```
#1) rede de palavras para quantificar e recortar uma amostra representativa;
#texto1 = "bola sapo prato caco"
#texto2 = "sapo caco frita brita"
#texto3 = "cc frita bota embrochavel"
#conjunto = c(texto1, texto2, texto2)
#tokens = tokens(conjunto)
\#tokens_ngrams(tokens, n = 3)
#2) correlação de palavras (n-grams);
#https://tutorials.quanteda.io/basic-operations/tokens/tokens_ngrams/
#https://quanteda.io/reference/tokens_ngrams.html
#n= integer vector specifying the number of elements to be concatenated in each ngram.
#Each element of this vector will define a n in the n-gram(s) that are produced
#skip = integer vector specifying the adjacency skip size for tokens forming the ngrams,
#default is 0 for only immediately neighbouring words. For skipgrams, skip can be a vector of integers,
#as the "classic" approach to forming skip-grams is to set skip = k where k is the distance for which k or fewer
#skips are used to construct the n-gram. Thus a "4-skip-n-gram" defined as skip = 0:4
#produces results that include 4 skips, 3 skips, 2 skips, 1 skip, and 0 skips
#(where 0 skips are typical n-grams formed from adjacent words). See Guthrie et al (2006).
#concatenator=character for combining words, default is _ (underscore) character
toks ngram <- tokens ngrams(v 0102.tokens, n = 2:3)
View(toks_ngram[1])
View(toks_ngram[[1]])
View(length(toks_ngram[[1]]))
Tokens consisting of 1 document.
     text1:
      [1] "homem_negro"
[4] "anos_espancado"
                                    "negro 60"
                                                             "60 anos"
                                    "espancado_morte"
                                                             "morte_assassinos"
      [7] "assassinos_seguranças" "seguranças_carrefour"
                                                            "carrefour_porto"
     [10] "porto_alegre"
                                   "alegre_entenderam"
                                                             "entenderam_homem"
     [ ... and 25 more ]
     'homem_negro' · 'negro_60' · '60_anos' · 'anos_espancado' · 'espancado_morte' · 'morte_assassinos' ·
     'assassinos_seguranças' · 'seguranças_carrefour' · 'carrefour_porto' · 'porto_alegre'
     'alegre_entenderam' · 'entenderam_homem' · 'homem_negro' · 'negro_60' · '60_anos' ·
     'anos assassinado' · 'assassinado criminosos' · 'criminosos trabalham' · 'trabalham carrefour'
     'homem negro 60' · 'negro 60 anos' · '60 anos espancado' · 'anos espancado morte'
     'espancado_morte_assassinos' · 'morte_assassinos_seguranças' · 'assassinos_seguranças_carrefour' ·
     'seguranças_carrefour_porto' · 'carrefour_porto_alegre' · 'porto_alegre_entenderam'
#2.1) filtrar correlação de palavras.
#Para filtrar só utilizar o parâmetro pattern
tokens_compound <- tokens_compound(v_0102.tokens, pattern = phrase("homem *"))</pre>
head(tokens compound, 1)
tokens_compound_select <- tokens_select(tokens_compound, pattern = phrase("homem*"))</pre>
head(tokens_compound_select[[1]], 30)
Tokens consisting of 1 document.
     text1:
      [1] "homem_negro" "60"
                                        "anos"
                                                       "espancado"
                                                                      "morte"
      [6] "assassinos"
                       "seguranças" "carrefour"
                                                                     "alegre'
                                                       "porto"
     [11] "entenderam" "homem_negro"
      ... and 6 more ]
     'homem_negro' · 'homem_negro'
#dfm_select remove features do dfm
#Como já retiramos anteriormente, não seria necessário refazer.
dfmat_inaug_nostop <- dfm_select(bag_of_words, pattern = stop_words, selection = "remove")</pre>
print(length(bag_of_words))
print(length(dfmat_inaug_nostop))
    [1] 52509389940
     [1] 52509389940
# dfm_remove tem o mesmo efeito que dfm_select remove
dfmat_inaug_nostop <- dfm_remove(bag_of_words, pattern = stop_words)</pre>
print(length(bag_of_words))
print(length(dfmat_inaug_nostop))
    [1] 52509389940
     [1] 52509389940
```

```
#You can also select features based on the length of features. In the example below, we only keep features consisting of at least five (
dfmat_inaug_long <- dfm_keep(bag_of_words, min_nchar = 5)</pre>
print(length(bag_of_words))
print(length(dfmat_inaug_long))
print(dfmat_inaug_long)
    [1] 52509389940
     [1] 46932121404
     Document-feature matrix of: 1,648,129 documents, 28,476 features (>99.99% sparse) and 0 docvars.
              homem negro espancado morte assassinos seguranças carrefour
       text1
                  0
                        1
                                   0
                                          0
                                                      0
                                                                 0
                                                                            0
                                                                                   0
       text2
                        0
                                         0
                                                     0
                                                                 0
                                                                                  0
       text3
                  0
                                   0
                                                                            0
                                         0
                                                                 0
                                                                            0
       text4
                  0
                        0
                                   0
                                                     0
                                                                                  0
       text5
                  0
                        0
                                   1
                                         0
                                                     0
                                                                 0
                                                                            0
                                                                                  0
       text6
                  a
                        a
                                                      a
                                                                 a
                                                                            a
                                                                                   a
             features
     docs
              alegre entenderam
       text1
                   0
                               0
       text2
                               0
                   0
                   0
                               0
       text3
       text4
                               0
                               0
       text5
       text6
     [ reached max_ndoc ... 1,648,123 more documents, reached max_nfeat ... 28,466 more features ]
topfeatures(bag_of_words, 50)
→ carrefour:
                   78048 é:
                                 49971 negro:
                                                   40415 homem:
                                                                      34164 assassinado:
                                                                         20458 negra:
           24700 espancado:
                                 23586 seguranças:
                                                       23563 morte:
           19891 dia:
                          18892 consciência:
                                                17226 anos:
                                                                  16904 #vidasnegrasimportam:
           16522 porto:
                           15389 alegre:
                                              15356 joão:
                                                               14469 assassinos:
                                                                                     13960 60:
           12594 racismo:
                              11129 brasil:
                                               11042 alberto:
                                                                  10104 preto:
                         8176 cada:
                                                            6905 caso:
                                                                           6809 supermercado
                                         7270 empresa
           8244 ser
#Revisar comando
#considerar esta célula para fazer o recorte da(s) frequência
#https://quanteda.io/reference/dfm_trim.html
#If max_docfreq = 0.1, features that occur in more than 10% of the documents are removed.
dfmat_inaug_docfreq <- dfm_trim(bag_of_words, max_docfreq = 0.5, docfreq_type = "quantile")</pre>
#print(length(featnames(bag_of_words)))
#print(length(featnames(dfmat_inaug_docfreq)))
#print(bag_of_words)
print(dfmat_inaug_docfreq)
Document-feature matrix of: 1,648,129 documents, 19,985 features (>99.99% sparse) and 0 docvars.
             features
     docs
              arrebentar tido normalizada <a href="https://t.co/hdu2koub3e">https://t.co/hdu2koub3e</a>
                       0
       text1
       text2
                       0
                             0
                                          0
                                                                   0
       text3
                       0
                             0
                                          0
                                                                   0
       text4
                       a
                             0
                                         a
                                                                   a
       text5
                       a
                             a
                                         0
                                                                   0
                                          0
       text6
              https://t.co/levm1m2mkv @fefeziiiiinha https://t.co/h03yfqudqx encerrei
       text1
                                                      0
                                                                                         0
       text2
                                                                                         0
                                     0
                                                      0
                                     0
                                                      0
                                                                               0
                                                                                         0
       text3
       text4
                                     0
                                                      0
                                                                               0
                                                                                         0
       text5
                                     0
                                                      0
                                                                               0
                                                                                         0
       text6
             features
     docs
              aspirador
                        edredom
       text1
                      0
       text2
                      0
                               0
       text3
                      0
                               0
       text4
                               0
       text5
                      0
       text6
     [ reached max ndoc ... 1,648,123 more documents, reached max nfeat ... 19,975 more features ]
```

Imagem sobre quantile:

https://miro.medium.com/max/546/1\*dkDFGXLRP64LtA26IqlsJg.png

Ngrams Vs Frequência Bag of Words

carrefour: 78048

```
tokens_compound_carrefour <- tokens_compound(v_0102.tokens, pattern = phrase("carrefour *"))
#Testar se funciona:
tokens_compound_select_carrefour <- tokens_select(tokens_compound, pattern = phrase("*carrefour*"))
tokens_compound_select_carrefour_inicio_da_palavra <- tokens_select(tokens_compound, pattern = phrase("carrefour_*"))
tokens_compound_select_carrefour_fim_da_palavra <- tokens_select(tokens_compound, pattern = phrase("*_carrefour"))</pre>
```