

# Acoso sexual callejero. Análisis cualitativo usando R

Elina Gómez

Agosto, 2018

## Contexto

El presente análisis se desarrolla en el marco del **estudio diagnóstico sobre el acoso sexual y otros tipos de violencia contra las mujeres y niñas en espacios públicos** en Montevideo. El mismo se enmarca en un convenio suscripto entre ONU Mujeres, la Asesoría para Igualdad de Género de la Intendencia de Montevideo (IM) y la Facultad de Ciencias Sociales (UdelaR).

## Definiciones

El estudio concibe al Acoso Sexual Callejero (ASC) como:

*“Incluye comentarios, atención, acciones o gestos con contenido sexual indeseado. Como en otras formas de violencia sexual, el componente clave del acoso sexual es que alguien realiza una acción sin el consentimiento, permiso o acuerdo de la persona o personas a las que se dirige. Acoso sexual incluye formas sin contacto, como comentarios sexuales acerca del cuerpo o apariencia de la persona, silbidos mientras pasa una mujer, exigencias de favores sexuales, quedarse mirando de manera sexualmente sugestiva, persecuciones y seguimiento de una persona, exposición de órganos sexuales. El acoso sexual también incluye formas físicas de contacto, como frotarse intencionalmente contra alguien en la calle o en el transporte público, agarrar, pegar una palmada y aproximarse a alguien en una manera sexual.” (ONU Mujeres, 2013)*

# Objetivos

- ▶ Realizar un análisis textual de grupos de discusión sobre ASC realizados a mujeres y varones.
- ▶ Analizar los datos surgidos de un Paper Review sobre la temática.

## I. Análisis textual

## Datos existentes

- ▶ 6 grupos de discusión sobre la temática ASC, en los cuales participaron mujeres y varones de Montevideo, pertenecientes a diferentes tramos etarios:

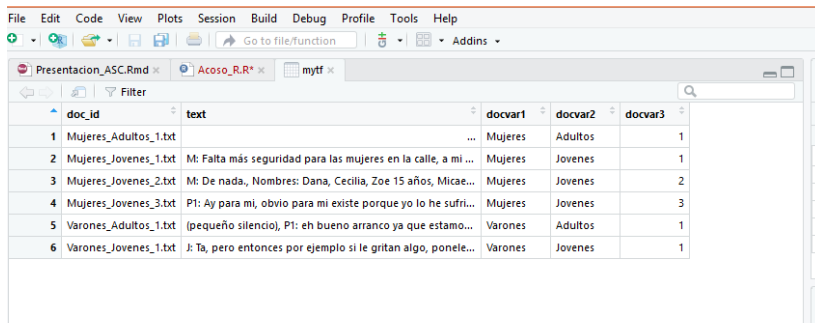
	Mujeres	Varones	Total
Jovenes	3	1	4
Adultos	1	1	2
Total	4	2	6

# Pre-procesamiento

- ▶ Abro los archivos de texto en los cuales tengo las transcripciones de los grupos
- ▶ Elimino aquellas menciones que corresponden a lxs moderadorxs
- ▶ Guardo como archivos .txt, identificando cada archivo por su nombre según las características del grupo (variables de agrupación posteriores)

# Pre-procesamiento

- ▶ Abro los archivos con la función **readtext()**, especificando ubicación de los mismos y variables anexas en los títulos para identificar características del documento



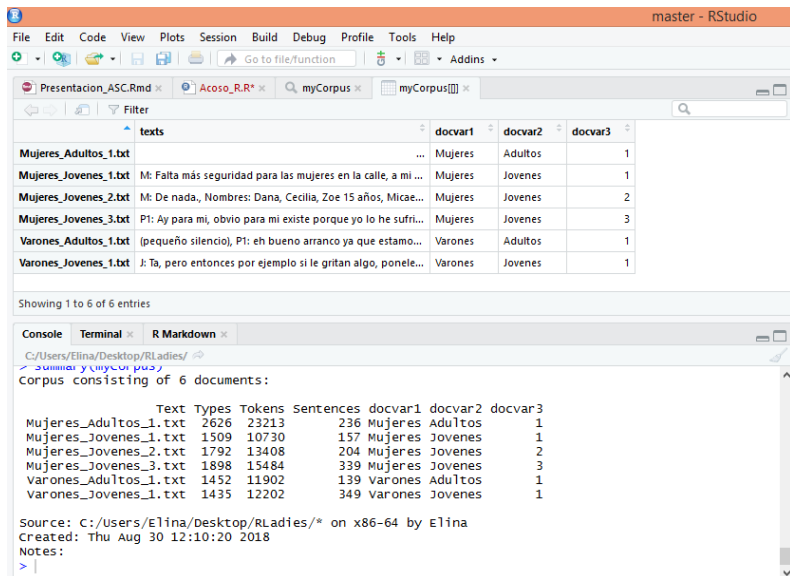
	doc_id	text	docvar1	docvar2	docvar3
1	Mujeres_Adultos_1.txt	...	Mujeres	Adultos	1
2	Mujeres_Jovenes_1.txt	Mi: Falta más seguridad para las mujeres en la calle, a mi ...	Mujeres	Jovenes	1
3	Mujeres_Jovenes_2.txt	Mi: De nada,, Nombres: Dana, Cecilia, Zoe 15 años, Micae...	Mujeres	Jovenes	2
4	Mujeres_Jovenes_3.txt	P1: Ay para mi, obvio para mi existe porque yo lo he sufri...	Mujeres	Jovenes	3
5	Varones_Adultos_1.txt	(pequeño silencio), P1: eh bueno arranco ya que estamo...	Varones	Adultos	1
6	Varones_Jovenes_1.txt	J: Ta, pero entonces por ejemplo si le gritan algo, ponele...	Varones	Jovenes	1

Figure 1



# Pre-procesamiento

- Creo un **corpus** con el paquete **quanteda**



The screenshot shows the RStudio interface with the 'myCorpus' object loaded in the Environment pane. The corpus contains 6 documents, each with a unique docvar1 and docvar2. The Console pane shows the output of the `summary(myCorpus)` command, displaying the number of tokens, sentences, and document variables for each document.

texts	docvar1	docvar2	docvar3
Mujeres_Adultos_1.txt	Mujeres	Adultos	1
Mujeres_Jovenes_1.txt	Mujeres	Jovenes	1
Mujeres_Jovenes_2.txt	Mujeres	Jovenes	2
Mujeres_Jovenes_3.txt	Mujeres	Jovenes	3
Varones_Adultos_1.txt	Varones	Adultos	1
Varones_Jovenes_1.txt	Varones	Jovenes	1

Showing 1 to 6 of 6 entries

```
Summary of myCorpus
corpus consisting of 6 documents:

      Text  Types Tokens Sentences docvar1 docvar2 docvar3
Mujeres_Adultos_1.txt 2626 23213      236 Mujeres Adultos      1
Mujeres_Jovenes_1.txt 1509 10730      157 Mujeres Jovenes      1
Mujeres_Jovenes_2.txt 1792 13408      204 Mujeres Jovenes      2
Mujeres_Jovenes_3.txt 1898 15484      339 Mujeres Jovenes      3
Varones_Adultos_1.txt 1452 11902      139 Varones Adultos      1
Varones_Jovenes_1.txt 1435 12202      349 Varones Jovenes      1

Source: C:/Users/Elina/Desktop/RLadies/* on x86-64 by Elina
Created: Thu Aug 30 12:10:20 2018
Notes:
>
```

## *“Limpieza” del texto*

- ▶ Creo un Document feature matrix (DFM), aplicando algunos argumentos que me permiten limpiar las palabras que no me interesan al efecto del análisis.
  - ▶ Homogeoizco las palabras en minúscula
  - ▶ Elimino números
  - ▶ Elimino puntuaciones
  - ▶ Elimino stopwords (por defecto y lista propia)
  - ▶ Elimino palabras con pocos caracteres (1 y 2)
  - ▶ Elimino palabras específicas que no me interesan (ej. “Risas”)

Nubes de palabras

## Nubes de palabras: general



## Nubes de palabras: grupos

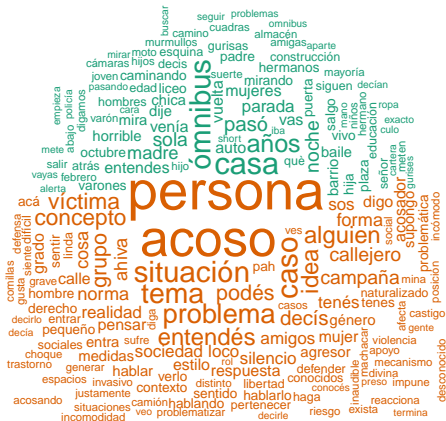
## Mujeres jóvenes



## Varones adultos

## Nubes de palabras: mujeres/varones

## Mujeres



## Varones

Nube de palabras: adultos/jovenes

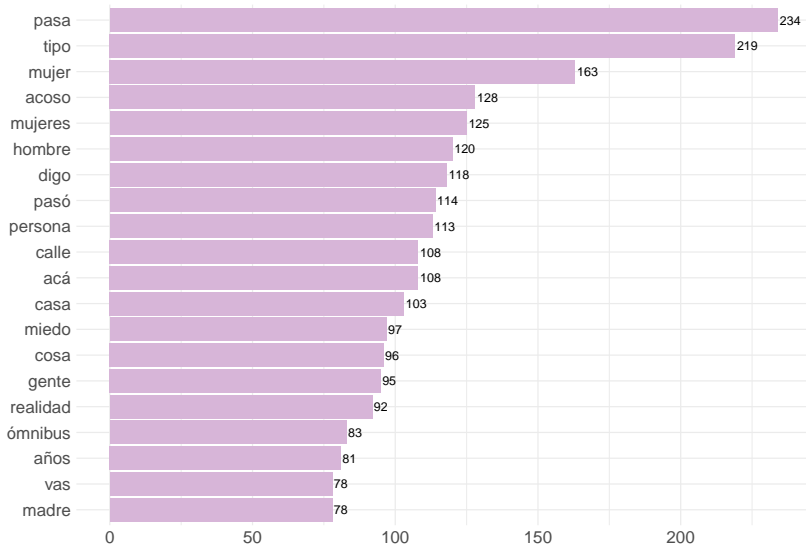
## Adultos



## Jovenes

# Palabras más frecuentes: general

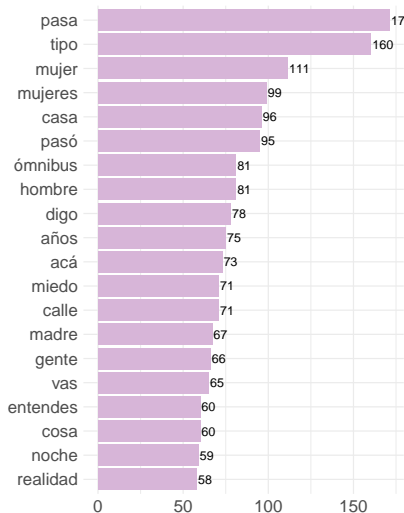
Palabras más frecuentes (n=15)



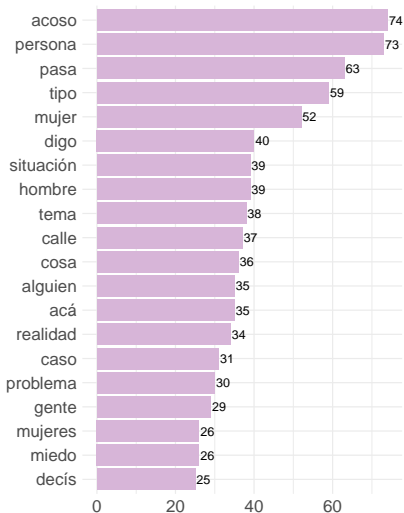


# Palabras más frecuentes: mujeres/varones

Palabras más frecuentes: Mujeres (n=20)



Palabras más frecuentes: Varones (n=20)

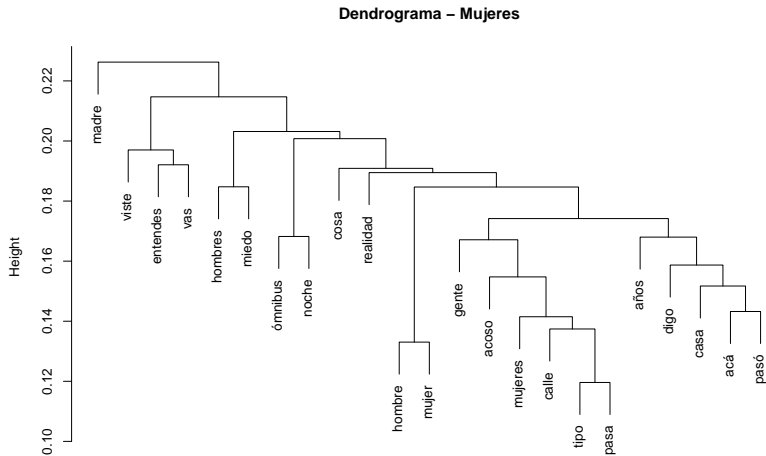


## Agrupación de palabras

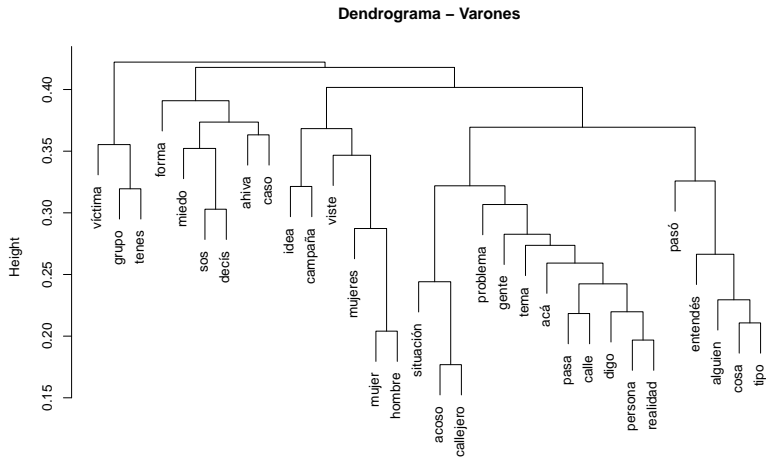
## Agrupación de palabras

- ▶ Para dar cuenta de la agrupación de palabras, uso el paquete **tm** y, a diferencia del análisis anterior, considero cada mención como un elemento del corpus
- ▶ Calculo las distancias entre los términos y armo *clusters* de palabras (método euclidiano)
- ▶ Grafico la agrupación jerárquica de las palabras en un dendograma, separada entre grupos (método de Ward)

# Agrupación de palabras: mujeres



# Agrupación de palabras: varones

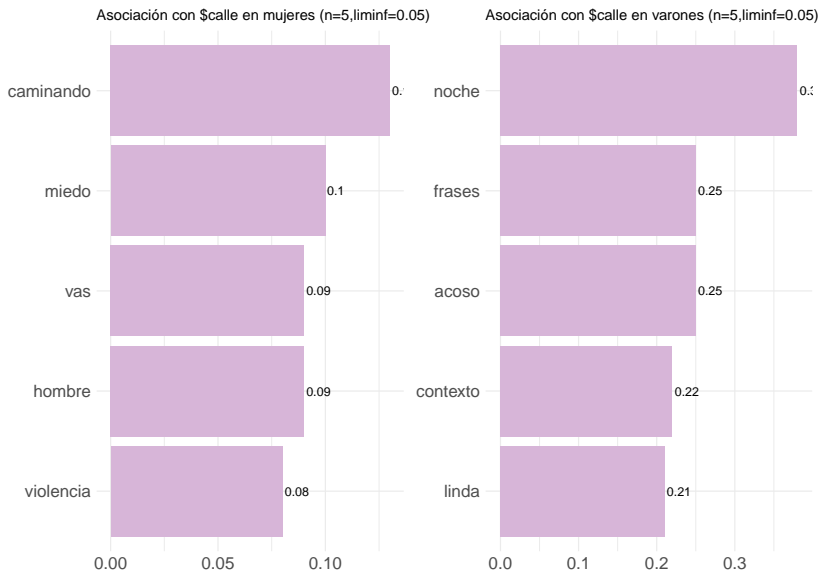


Asociación de palabras

# Asociación de palabras

- ▶ Para analizar las asociaciones entre palabras, consideramos una dispersión de 0.995
- ▶ Buscamos la asociación de palabras, considerando una correlación mínima de 0.05
- ▶ Al tratarse de menciones y no de grupos, la dispersión entre términos es muy alta
- ▶ Analizamos la asociación con las palabras: *calle*, *acoso* y *mujer*

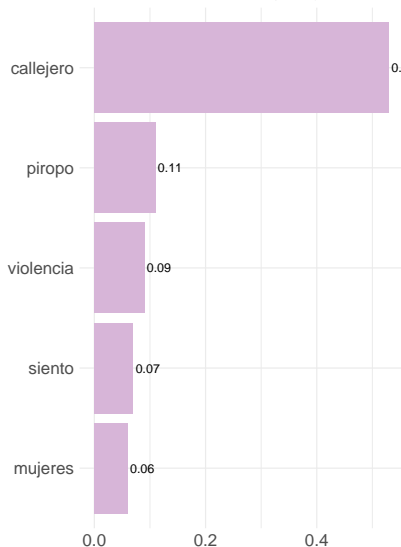
# Asociación de palabras: \$calle



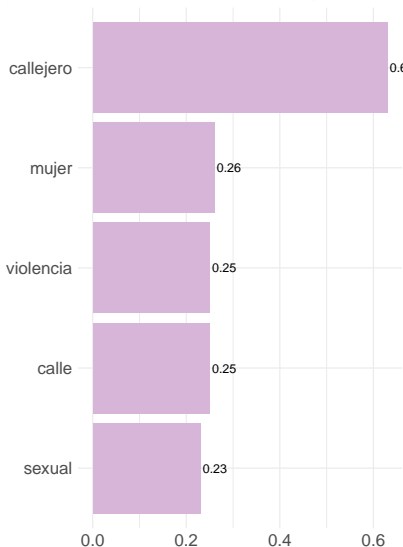


# Asociación de palabras: \$acoso

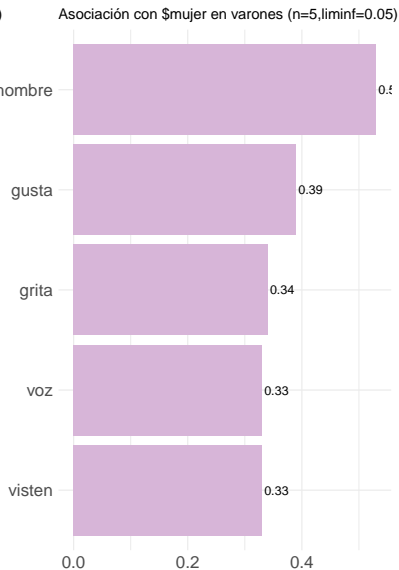
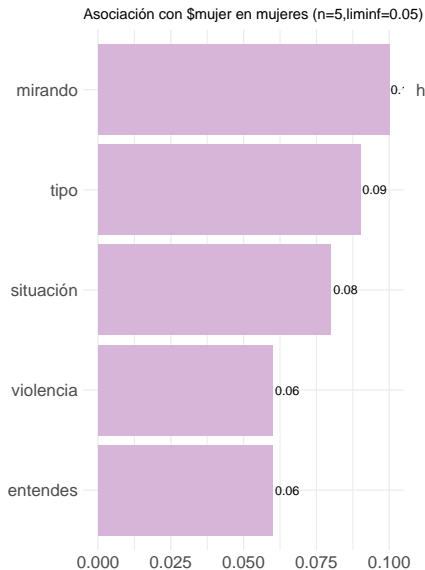
Asociación con \$acoso en mujeres (n=5, liminf=0.05)



Asociación con \$acoso en varones (n=5, liminf=0.0)



# Asociación de palabras: \$mujer



## Análisis de sentimiento

# Análisis de sentimiento

Se presentan dos métodos para analizar sentimiento de los documentos:

- ▶ Diccionario LWIC-Spanish: con la función `dfm_lookup()` de *quantda* identifica en los documentos las emociones presentes en el diccionario y establece puntajes para cada uno, a partir de la estandarización de los mismos (conicidencia únicamente de 15% entre términos).
- ▶ Método Syuzhet: utiliza la función `get_sentiment()` de *syuzhet* asigna puntajes a cada documento según el método y lenguaje indicado. El método **syuzhet** es un diccionario de sentimientos desarrollado en el Laboratorio Literario de Nebraska. Otros métodos: *bing*, *afinn*, *nrc* y *stanford*.

# Análisis de sentimiento

Comparación entre los métodos:

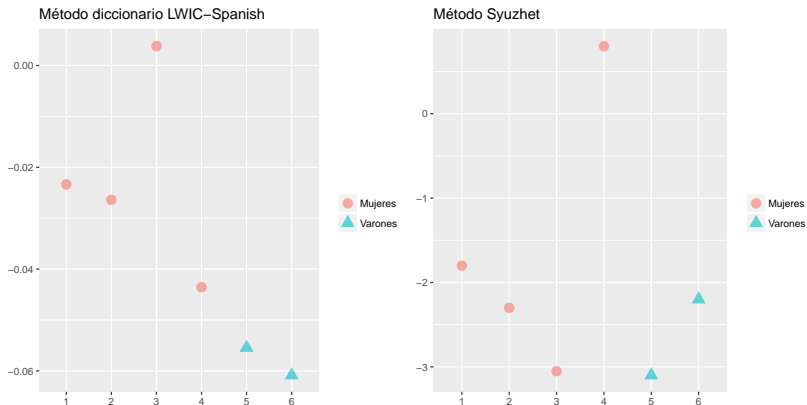


Figure 3

# Análisis de sentimiento

Sentimientos más extremos – Mujeres



# Análisis de sentimiento



## II. Paper review

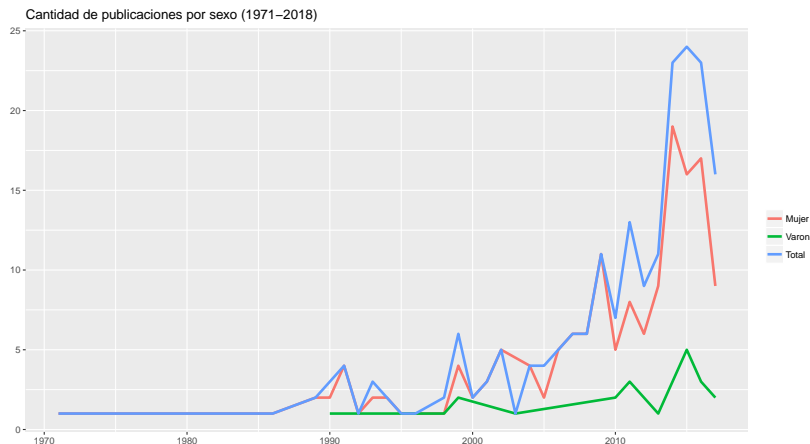


## Paper review

- ▶ A partir de la sistematización de los trabajos los diversos tipos de documentos hallados, se construyó un catálogo con 202 trabajos académicos de distinta índole, a saber: artículos arbitrados, ensayos, reportes, informes, monografías, ponencias, libros, que abarcan variedad de países y regiones dentro de los cinco continentes.

# Sexo

- ▶ Al separar por sexo la autoría de los trabajos, se destaca que las mujeres investigan casi cinco veces más que los varones en la temática

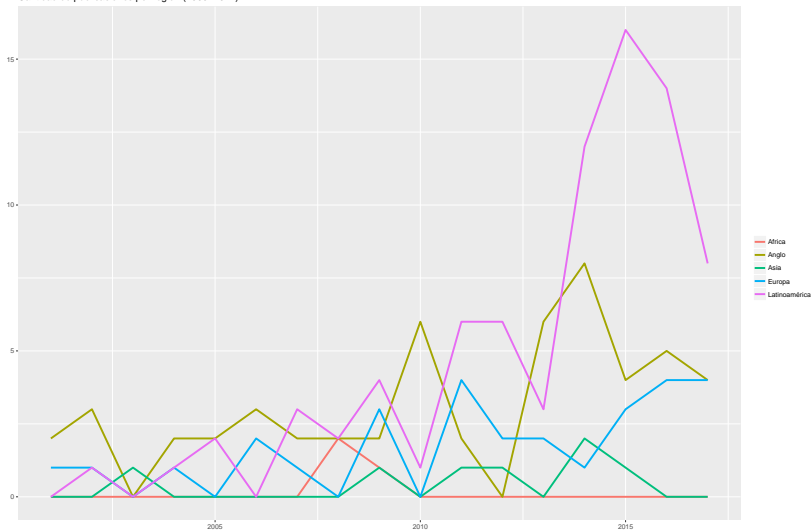


# Región

- ▶ De los trabajos relevados casi seis de cada diez son escritos en español, y de estos un 70 % han sido producidos en Latinoamérica.
- ▶ Antes de 2008 la producción era casi exclusivamente en inglés, y a partir de esa fecha se potencia la investigación y acumulación en América Latina.

# Región

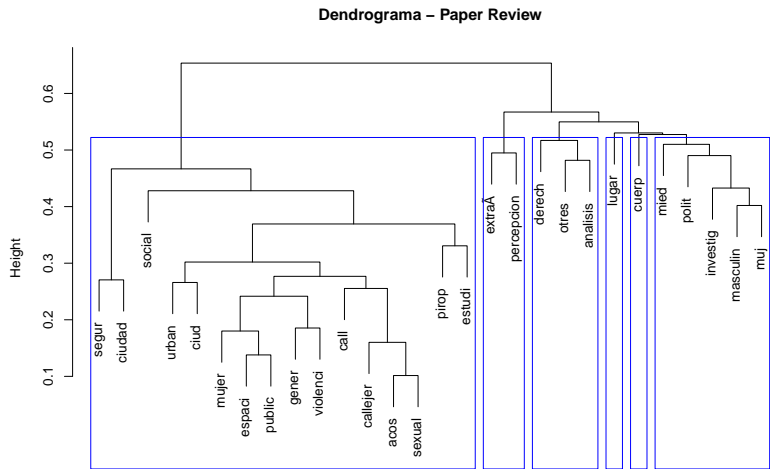
Cantidad de publicaciones por región (2000-2017)



## Análisis de títulos: décadas



# Dendrograma



Muchas gracias!!