



# Contando hombres y mujeres con Damegender GNU/Linux como caso de uso

David Arroyo Menéndez ([davidam@gnu.org](mailto:davidam@gnu.org))

Jesús González Barahona ([jgb@gsyc.es](mailto:jgb@gsyc.es))

Gregorio Robles ([grex@gsyc.es](mailto:grex@gsyc.es))




# Índice

- Marco teórico
- Fuentes secundarias sobre brecha de género (United Nations Women, Global Gender Gap Index 2012, ESA calculations, ...)
- Lo técnico de las cuentas: elegir dataset y decidir sobre el sexo del nombre
- Contando nombres y género en GNU/Linux como caso de uso



# Marcos Teóricos

Necesitamos los conceptos en los contextos



# Marco Teórico (0): Vamos a tratar de detectar los conceptos y las filosofías que vamos a aplicar

Tenemos que poner las palabras en los contextos para entender de qué estamos hablando, pero recordando a qué filosofías nos estábamos refiriendo.

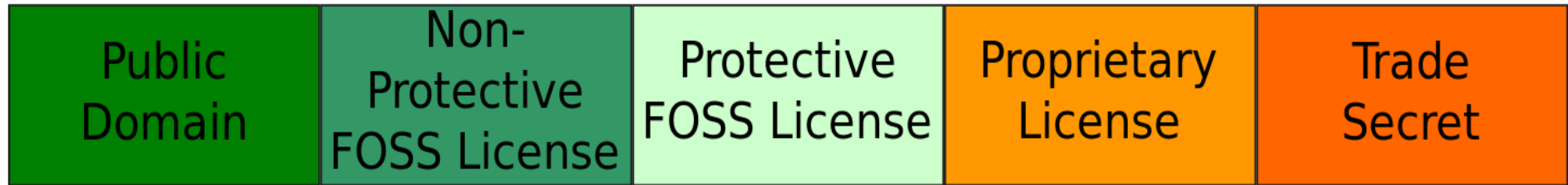
# Marco Teórico (I): En las notas de copyright hay valores

Ejemplo de una nota de copyright de software privativo sería algo como.

(c) Todos los derechos reservados. Amazon 2021

# Marco Teórico (II): Grados en el copyright que solemos encontrar

## Rights in Copyright



All rights  
relinquished

←  
more rights granted

→  
more rights retained

All rights  
retained

# Marco Teórico (III): ¿Es posible crear discriminación positiva de género en el copyright?

Podría escribirse una nota de copyright del estilo:

Se permite la copia, modificación y redistribución de las copias siempre y cuando usted sea una mujer.

# Marco Teórico (IV): En la idea de nombre y origen nacional hay valores

La multiculturalidad se refiere a la aceptación y el reconocimiento de las diferentes culturales desde el respeto.

La interculturalidad se refiere al enriquecimiento mutuo desde el cruce de esas culturas ya identificadas.



# Marco Teórico (V): Reconocer origen nacional versus origen internacional

Nombres como David, Sara, Ana, Andrea, ... tienen presencia en múltiples naciones

Detectar el origen nacional mejora la precisión del acierto de género

Ocultar el origen, anonimiza para evitar la represalia a las personas con un origen diferente al de la región.

Hemos generado un conjunto de datasets con licencias libres de múltiples fuentes estadísticas con los orígenes nacionales que cubre casi todo occidente y un dataset libre también de carácter internacional.

# Marco Teórico (VI): Feminismo, Ecofeminismo

El feminismo se refiere a la mejora de los derechos de las mujeres

La equidad de género se refiere a detectar donde hay brecha de género y solventarla

El ecofeminismo tiene una perspectiva en la que la lucha de las mujeres está correlacionada con la salvaguarda del planeta con una lógica de conservación y respeto a la vida en su conjunto.

# Marco Teórico (VI): Feminismo, Ecofeminismo

El feminismo se refiere a la mejora de los derechos de las mujeres

La equidad de género se refiere a detectar donde hay brecha de género y solventarla

El ecofeminismo tiene una perspectiva en la que la lucha de las mujeres está correlacionada con la salvaguarda del planeta con una lógica de conservación y respeto a la vida en su conjunto.

# Marco Teórico (VII): Interseccionalidad

La interseccionalidad se refiere al hecho de que no se pueden entender los valores de una manera aislada, sino que generalmente actúan como un todo de manera que si se elimina feminismo, esto le afecta a la lucha obrera, el ecologismo, etc.

# Marco Teórico (VIII): Género, sexo, ...

Refiriéndonos al tema de hombre o mujer, lo masculino respecto a lo femenino lo podemos denominar sexo y generalmente nos referimos al sexo biológico con lo que generalmente se nace o se opera.

También lo podemos denominar género que tiene que ver con como uno se identifica, o se siente.

Muchas personas se identifican como no binarias debido a que este constructo social cambia con los afectos y la cultura política sexual en la que uno está inmerso.

Las legislaciones en los estados solo han traído datasets binarios, por lo que hemos utilizado porcentajes y frecuencias para entender masculinidades y feminidades en los nombres.

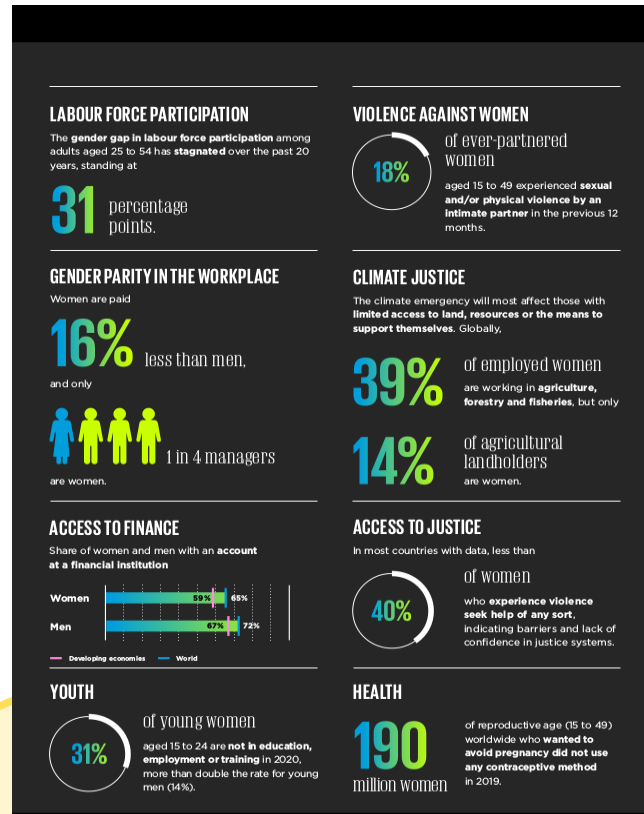


# Fuentes Documentales

Necesitamos comparar nuestros datos con otros estudios

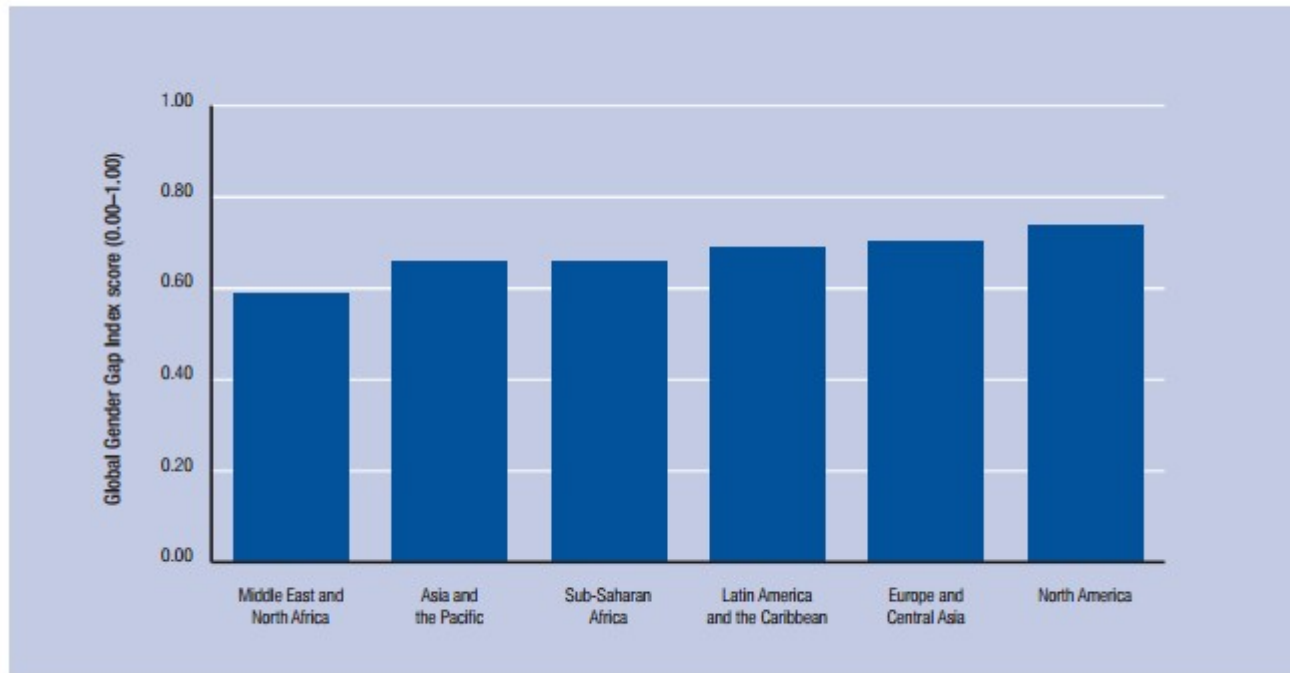


# Fuentes Secundarias (I): Brecha de Género en términos generales



# Fuentes Secundarias (II): Brecha de Género en términos generales

Figure 2: Regional performance on the Global Gender Gap Index 2012

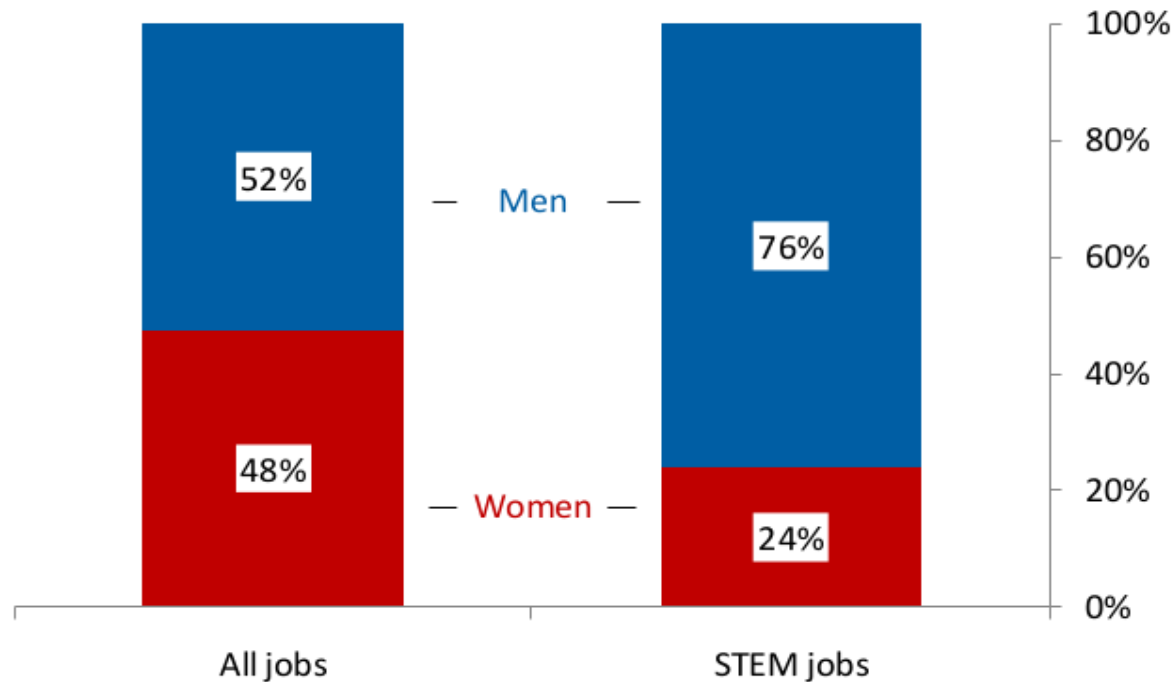


Source: Global Gender Gap Index 2012; details of regional classifications in Appendix B.  
Scores are weighted by population; population data from the World Bank's *World Development Indicators (WDI)* online database 2011, accessed July 2012.



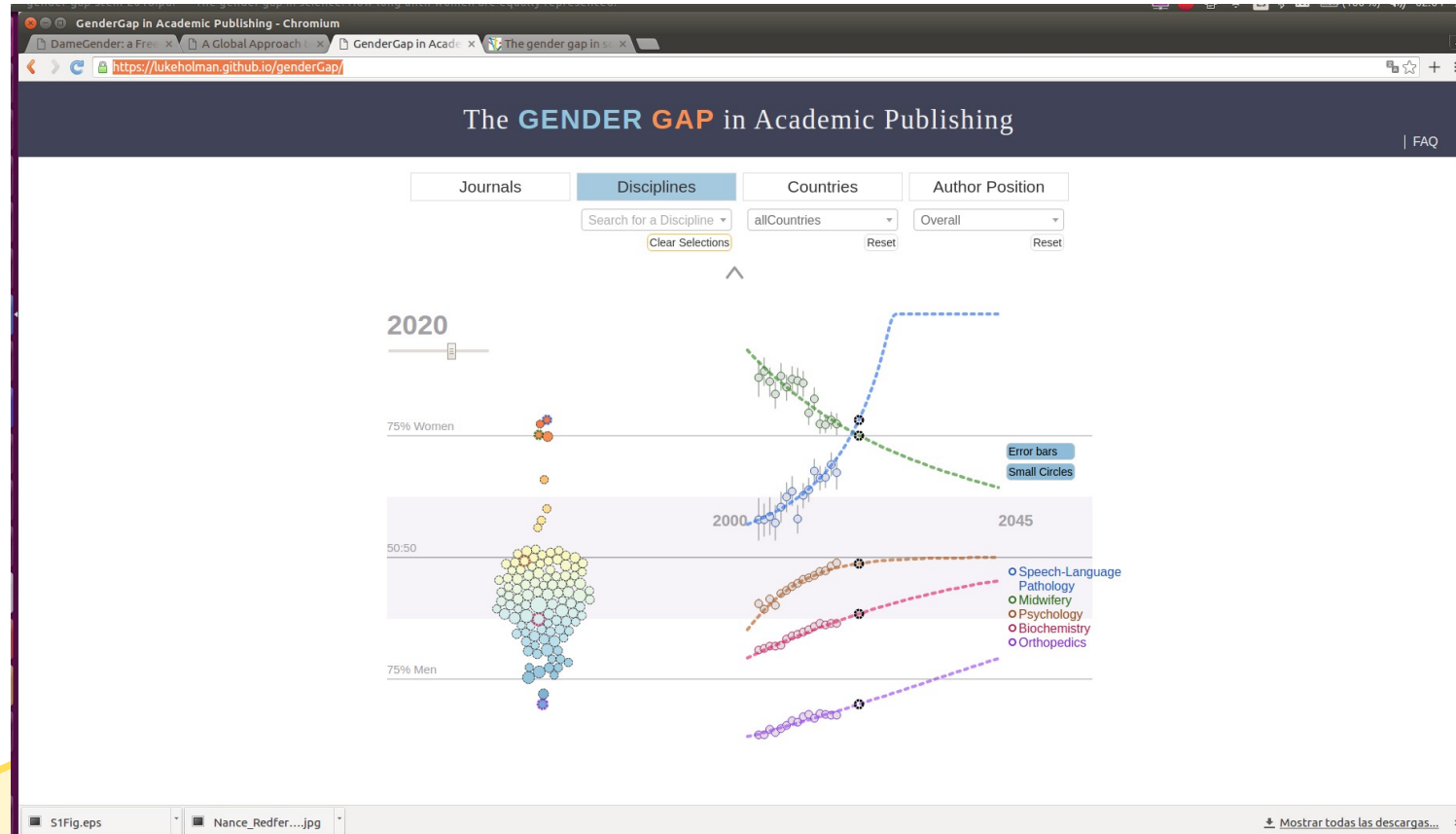
# Fuentes Secundarias (III): Brecha de Género en STEM

**Figure 1. Gender Shares of Total and STEM Jobs, 2009**



Source: ESA calculations from American Community Survey public-use microdata.  
Note: Estimates are for employed persons age 16 and over.


# Fuentes Secundarias: Luke Holman





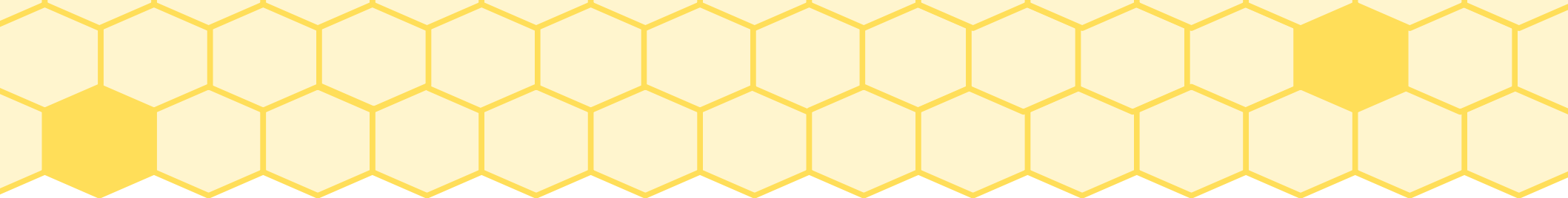
# Datasets en Damegender

21 datasets open data nacionales  
1 dataset internacional desde estos datasets




# Datasets en Damegender






Cuando el nombre no está en el dataset  
(machine learning o pagar a la api)





# Indicadores de calidad en el acierto con los nombres

accuracy, precision, f1-score, confusion matrix



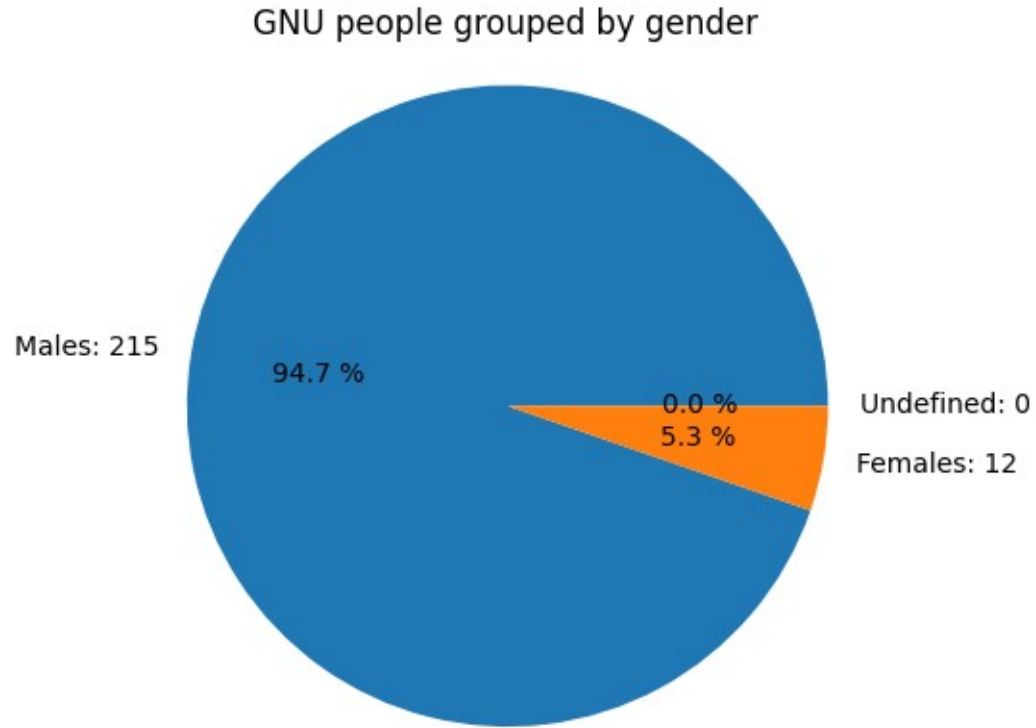


# Midiendo Desigualdad de Género en GNU/Linux con Damegender

Estos son nuestros datos preliminares

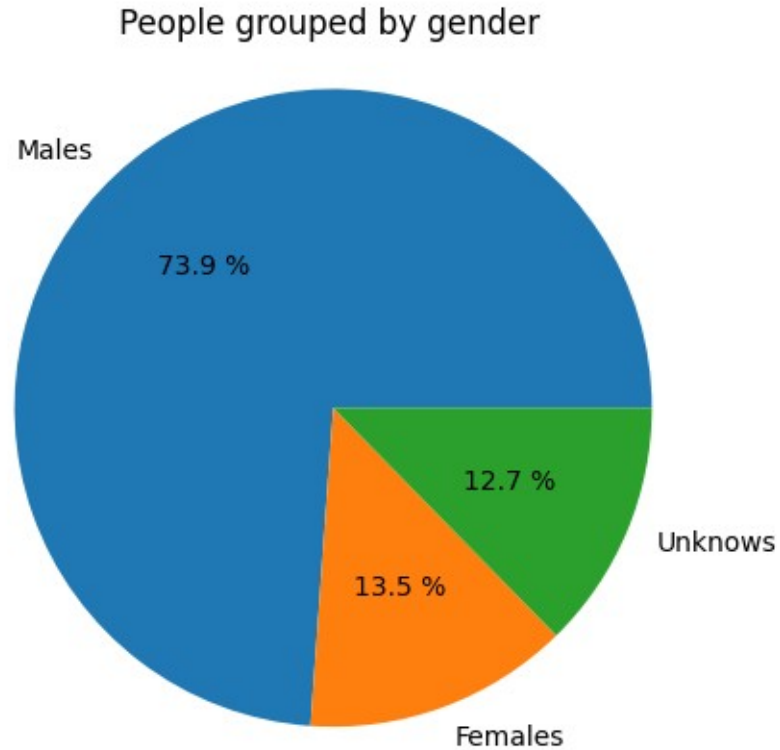


# Midiendo con Damegender (I): GNU

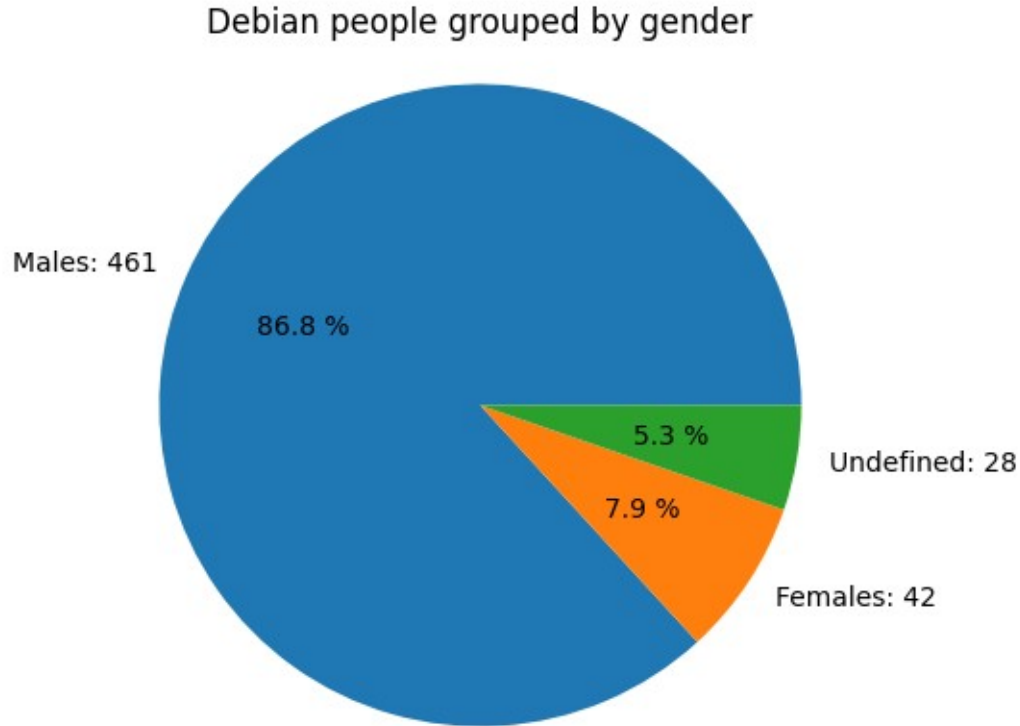


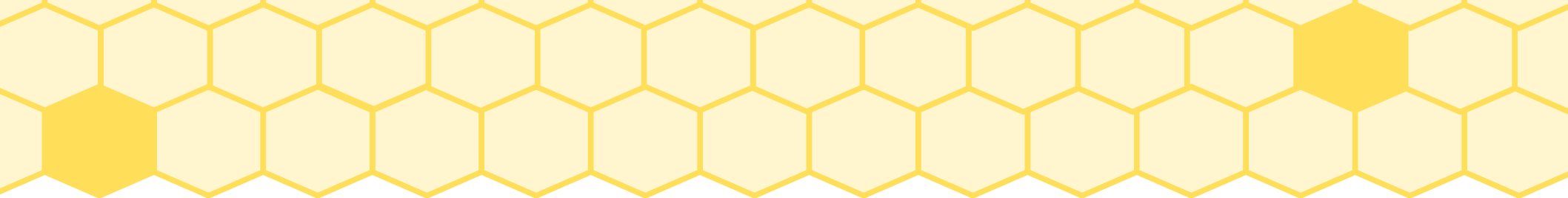


# Midiendo con Damegender (II): Linux



# Midiendo con Damegender (III): Debian



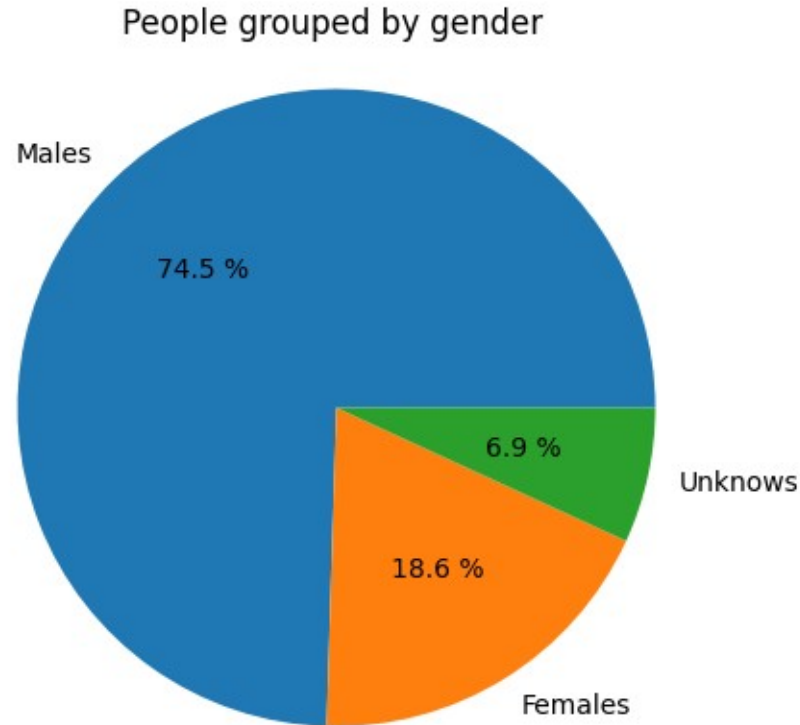


# Midiendo Desigualdad de Género con Damegender en otros datasets

Estos son nuestros datos preliminares

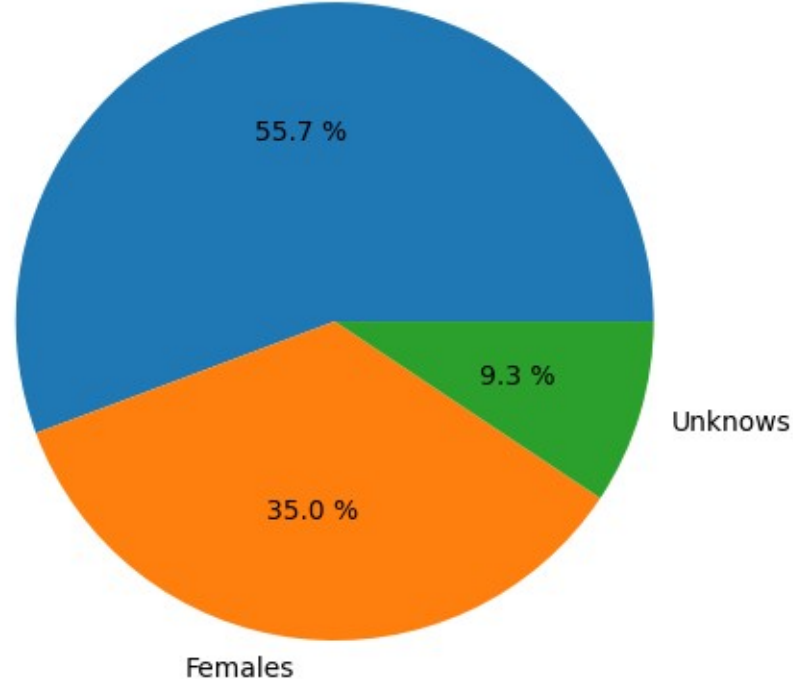


# Midiendo con Damegender (IV): 119 de Forbes



# Midiendo con Damegender (V): Webometrics

Scientific people grouped by gender  
Males



# Conclusiones y Trabajos Futuros

- Existe desigualdad de género
- Es posible medirla con datasets internacionales libres