



**WOMEN IN DATA SCIENCE**  
MONTEVIDEO

MUJERES LÍDERES EN CIENCIA DE DATOS - TÉCNICAS Y  
EXPERIENCIAS



WOMEN IN DATA SCIENCE



Plan Ceibal



# El discurso de género en el Parlamento uruguayo

**Elina Gómez Bonaglia**

Socióloga | Unidad de Métodos y Acceso a Datos (UMAD) / Udelar



WOMEN IN DATA SCIENCE



Plan Ceibal

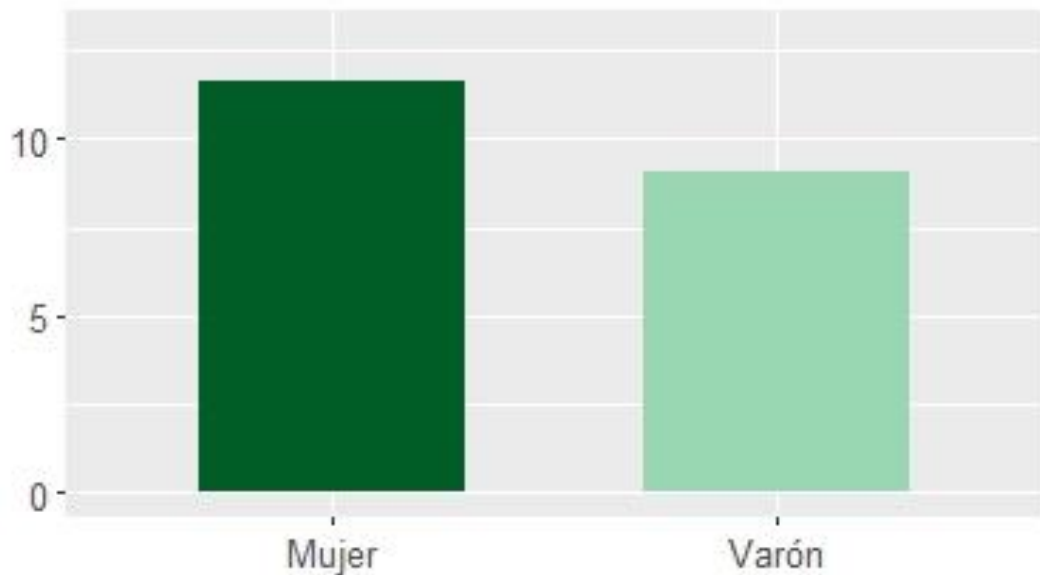
# OBJETIVO GENERAL

El **objetivo general** del trabajo es lograr un buen método de clasificación de texto que permita distinguir entre las intervenciones parlamentarias en la Cámara de representantes (unidad mínima de análisis) que planteen y discutan la temática de género para luego analizar mediante variables anexas, qué representantes, partidos, sectores son los que instalan la discusión en dicha materia.

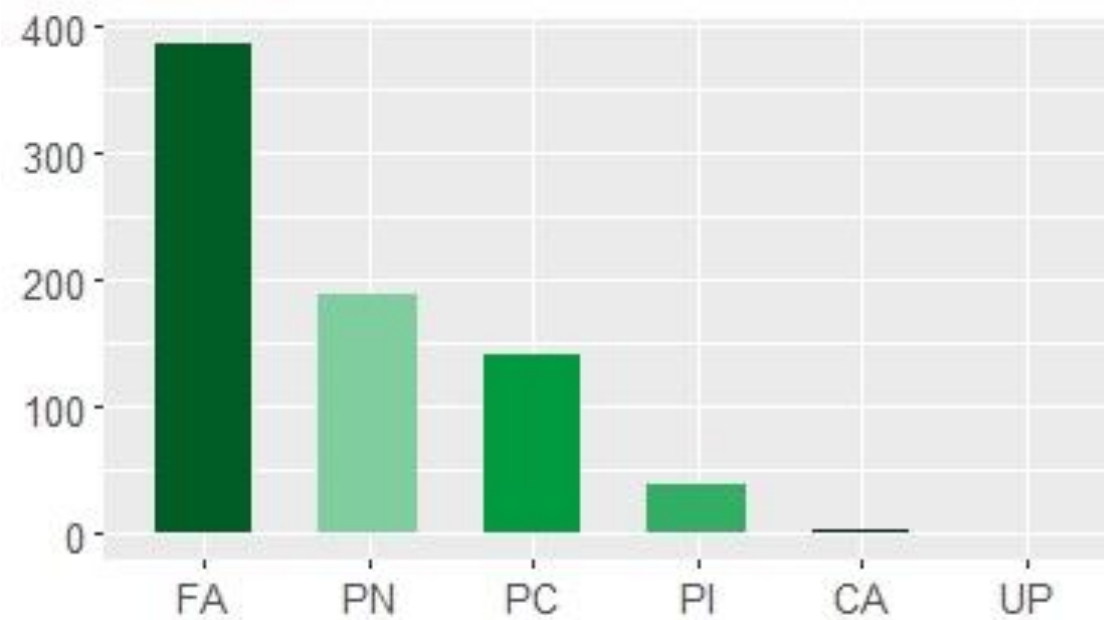


# ALGUNOS RESULTADOS

Peso relativo de menciones por sexo

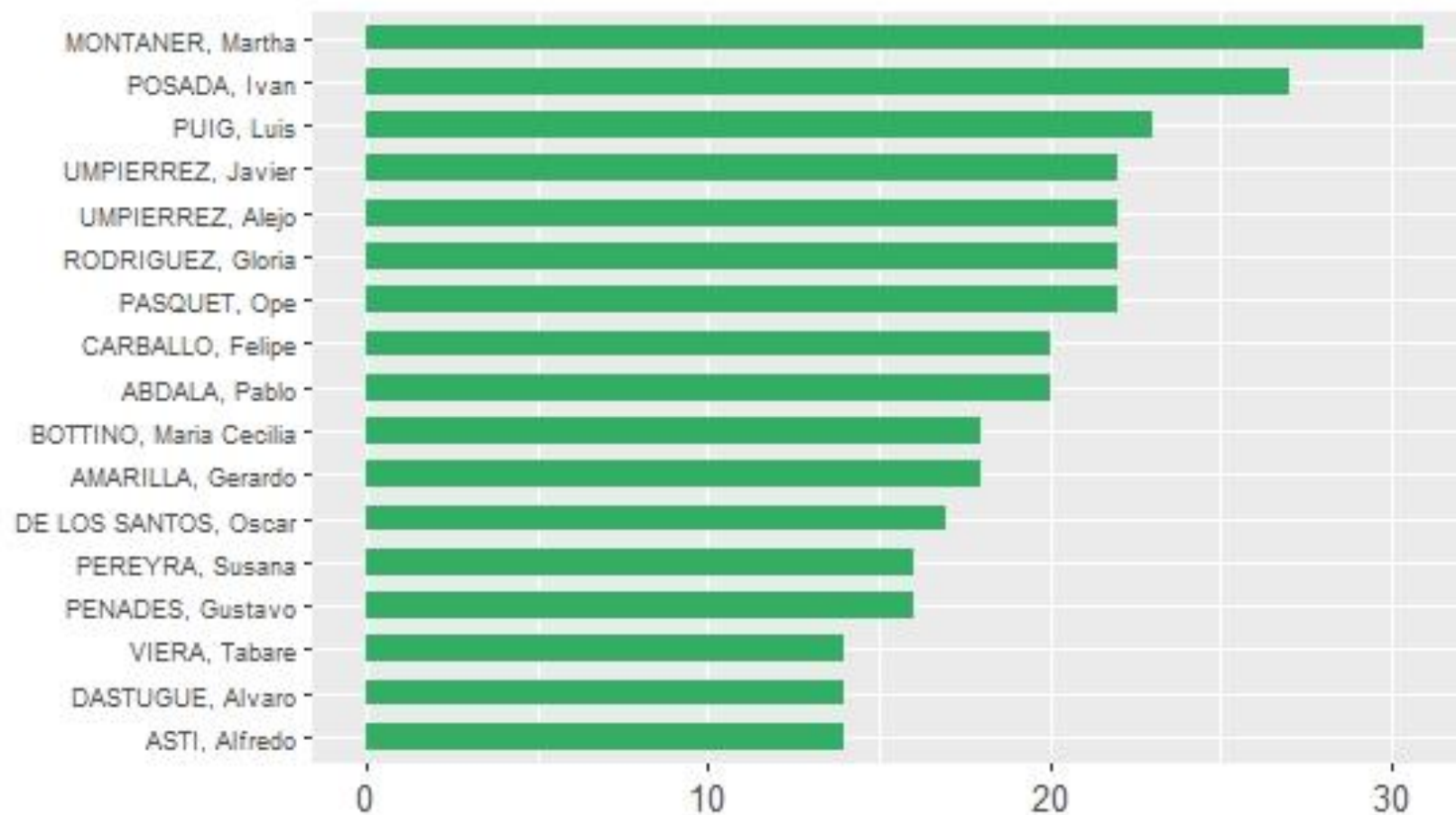


Cantidad de menciones por partido político



# ALGUNOS RESULTADOS

Cantidad de menciones por legislador/a



# REQUISITOS PREVIOS

1. Instalar R y RStudio
2. Descargar materiales necesarios disponibles en:  
[https://github.com/elinagomez/parlamentoygenero\\_Wids2022](https://github.com/elinagomez/parlamentoygenero_Wids2022)
3. Instalar paquetes necesarios que aparecen en el script



# PASOS A SEGUIR

1. Obtención de la información
2. Limpieza del texto y matriz de términos
3. Clasificación: aprendizaje automático (ML) y diccionario
4. Análisis de los datos



# 1. Obtención de la información

- Para la obtención de información, usaremos el paquete speech el cual permite descargar las sesiones parlamentarias de forma fácil y ordenada, agregando a los discursos algunas variables anexas como la legislatura, el nombre, la fecha de la sesión, entre otras.
- Para agregar la etiqueta partidaria usamos el paquete puy que cuenta con una base histórica de político/as uruguayos/as.





# 1. Obtención de la información

- Otra forma accesible sería descargar la base de forma manual con la aplicación:

speech App →



# 1. Obtención de la información

- Para este taller vamos a utilizar la base de sesiones ya descargada correspondiente a la la legislatura pasada: XLVIII (2015-2020), y a la cámara de Diputados/as, con etiqueta partidaria incorporada.
- Para entrenar el modelo, usaremos una base con intervenciones de la Comisión de Género del parlamento y otras Comisiones temáticas.



- 1. Clasificación: aprendizaje automático (ML) y diccionario**
- 2. Análisis de los datos**



# ¡Muchas gracias!



Plan Ceibal



WOMEN IN DATA SCIENCE