



CURSO ONLINE **Ecofemidata**

CLASE 1: Introducción a los datos con perspectiva feminista

DOCENTES: Celina Santellan y Carolina Pradier

Organización del curso

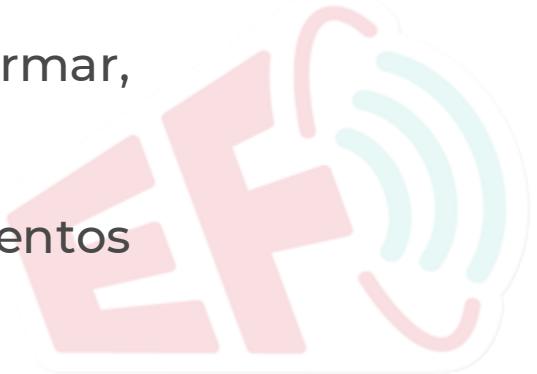
- Las clases se llevarán a cabo de manera **asincrónica** cada semana. Constarán de una parte **teórica**, guiada por un **video** y un **pdf** y una parte donde pondrán en **práctica** los conocimientos aprendidos.
- Lxs alumnxs tendrán la oportunidad de presentar dudas en el **foro** del campus y en una clase **sincrónica** cada dos semanas.



Objetivos

El curso propone que lxs alumnxs puedan:

- a. Introducir y desarrollar los fundamentos conceptuales, teóricos, procedimentales, metodológicos y prácticos para el manejo del programa Rstudio.
- b. Indagar en las funciones de los paquetes *tidyverse* y *ggplot* para lograr una práctica en el manejo de los datos e información estadística.
- c. Desarrollar capacidades para importar, ordenar, transformar, visualizar, y comunicar datos.
- d. Diseñar e implementar un informe que articule los conocimientos adquiridos dentro del espacio.



¿Qué es R?

- R es un lenguaje de programación especializado en **análisis y visualización** de datos.
- Es un producto de código abierto! Cualquier persona puede usarlo y modificarlo sin pagar licencias ni costos de adquisición.
- Además del lenguaje en sí (**R base** - sintaxis básica), se pueden incorporar cientos de librerías (o **paquetes**) con distintas funciones desarrolladas por los usuarios.
- Existe una comunidad muy grande para realizar preguntas y despejar dudas (aprender cómo googlear!).



Crecimiento de la cantidad de datos y la capacidad para procesarlos

- Crecimiento exponencial de la cantidad de datos que se recolectan y almacenan, así como también de las capacidades para procesarlos.
- Distinguir los **datos** (información de que algo sucedió) de los **conocimientos** (información acerca de por qué algo sucedió).
- Nosotras vamos a usar R para explotar datos y extraer conocimientos.



Datos con perspectiva feminista



Mostrar que los problemas son sistémicos.

“Lo personal es político”.

¿Por qué nos interesa el análisis cuantitativo?

Al traducir experiencias similares experimentadas por grandes grupos de personas a datos, podemos encontrar patrones, demostrar que los problemas son sistémicos, podemos proponer mejores soluciones.



Mostrar que los problemas son sistémicos.

“Lo personal es político”.

No siempre es suficiente con datos!

Se ha estudiado que aquellas personas en una situación privilegiada creen sin cuestionar evidencia anecdótica de sus pares, pero ante las personas pertenecientes a minorías, exigen evidencia exhaustiva -y cuestionan su validez- (para profundizar, ver el trabajo de Candice Lanius).



Identificar qué **potencialidades y **limitaciones** tienen los datos**

- ¿Qué falta?
- ¿Qué datos no se miden?
- ¿Qué se puede lograr con los datos que sí tenemos?



Identificar qué **potencialidades** y **limitaciones** tienen los datos

Algunos casos:

- Las encuestas de uso del tiempo que permitieron cuantificar la desigual distribución de trabajo no remunerado al interior de los hogares empezaron a relevarse mucho después que las que miden las horas de trabajo remunerado.
- [Feminicidios en México](#): mujeres toman la tarea que el Estado no lleva a cabo.
- Una iniciativa para darle visibilidad a estos problemas: [The Library of Missing Datasets](#).



Identificar qué **potencialidades** y **limitaciones** tienen los datos

¿Qué determina que algunos datos se recolecten y otros no?

- Recolectar y almacenar datos es costoso (energía!).
- Hay información difícil -no imposible!- de cuantificar (especialmente en torno a lo que sucede en la esfera privada -ej: ¿cuántas mujeres sufren violencia psicológica?-), y hay información que no se considera importante -ej: raza en los estudios de mercado de trabajo en Argentina-.
- ¿Quién diseña los cuestionarios que determinan los datos de las encuestas y censos? ¿Quién elige las variables de interés acerca de patrones de consumo/movilidad/etc. que se guardan y analizan?



Identificar qué **potencialidades** y **limitaciones** tienen los datos

La paradoja de la exposición

A veces, hacer visibles para el sistema a algunos colectivos de personas -lo cual es necesario para obtener recursos para mejorar su calidad de vida- puede resultar en situaciones de violencia/discriminación.

Si el colectivo de personas no es grande, ¿los datos son desanonomizables?

El caso de los inmigrantes sin papeles:

Si un censo se identifica una zona con una población muy grande de este tipo, ¿van a llegar más trabajadores y trabajadoras sociales para ayudarles a integrarse? ¿o se van a redoblar los esfuerzos por encontrarlos, detenerlos y deportarlos?



¿Las prácticas actuales refuerzan las desigualdades existentes?

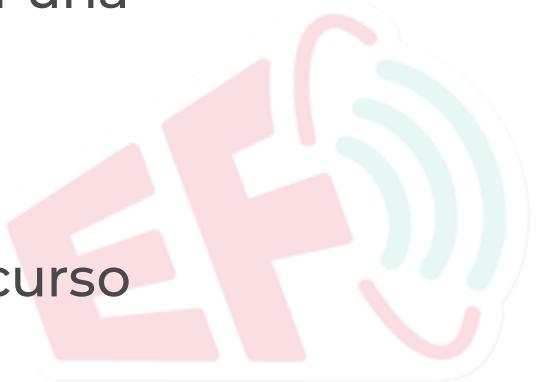
Ausencia de mujeres -y otras minorías- en STEM

La ciencia no es neutral, trae los prejuicios de sus científicos.

Las formas de trabajo y comunicación que pretenden asegurar objetividad usualmente presentan el punto de vista hegemónico.

Es preferible explicitar los contextos en los que se desarrolla una investigación, y las posibles fuentes de sesgo.

Para profundizar sobre estos temas, se puede consultar el curso de [Ciencia y Feminismo](#) que ofrece Ecofeminita.



¿Las prácticas actuales refuerzan las desigualdades existentes?

Ausencia de mujeres -y otras minorías- en STEM

Algunos ejemplos:

No es neutral que mucha de la investigación en salud y seguridad tome como modelo el cuerpo de un hombre blanco.

- Estadísticamente es más probable que un hombre esté involucrado en un accidente de tránsito. Pero si una mujer tiene un accidente, tiene 47% más chances de ser herida de gravedad y 17% más chances de morir que el hombre. ¿[Por qué](#)? Los crash test dummies fueron diseñados con la forma, tamaño y musculatura de un varón promedio.
- El caso de [Joy Buolamwini](#), científica afroamericana del MIT, que descubrió que el software de **reconocimiento facial** de un proyecto de investigación no reconocía su cara: había sido entrenado con un 78% de caras pertenecientes a varones y un 84% a personas blancas.



Etapas para *intervenir* sobre los sesgos y prejuicios de una investigación

1. La construcción de la base de datos

Para construir una base de datos, es necesario establecer variables (y sus categorías) para clasificar las observaciones (personas, hogares, empresas, etc.).

Al hacerlo, estamos construyendo una versión simplificada de la realidad donde vamos a representar al objeto de estudio de acuerdo con *algunas* de sus características.



Etapas para **intervenir** sobre los sesgos y prejuicios de una investigación

1. La construcción de la base de datos

¿Cuáles son las variables relevantes?

- Horas diarias de trabajo no remunerado?*
- Exigencia de horas extra?*
- Jornada laboral compatible con horarios escolares?*
- Cantidad de episodios de violencia en la empresa en el último año?*
- Cantidad de trabajadoras con cuerpos no hegemónicos en la empresa?*

¿Y las categorías relevantes?

- Mujer/Varón?*
- Pregunta PP02E de la EPH: Durante esos 30 días, no buscó trabajo porque... 1= está suspendido 2= ya tiene trabajo asegurado 3= se cansó de buscar trabajo 4= hay poco trabajo en esta época del año 5= por otras razones.*



Etapas para **intervenir** sobre los sesgos y prejuicios de una investigación

2. Análisis de los datos y comunicación de los resultados

- Definir las preguntas y dialogar con otras perspectivas

-¿a quién queremos ayudar? ¿podemos establecer un diálogo con ese colectivo?

-¿qué queremos preguntarle a los datos? ¿por qué las mujeres no estudian ingeniería o a qué situaciones de violencia de género se enfrenta una ingeniera o estudiante de ingeniería en su día a día?

-¿qué tipo de personas están interviniendo en la toma de decisiones? ¿son todas blancas y de clase alta?



Etapas para **intervenir** sobre los sesgos y prejuicios de una investigación

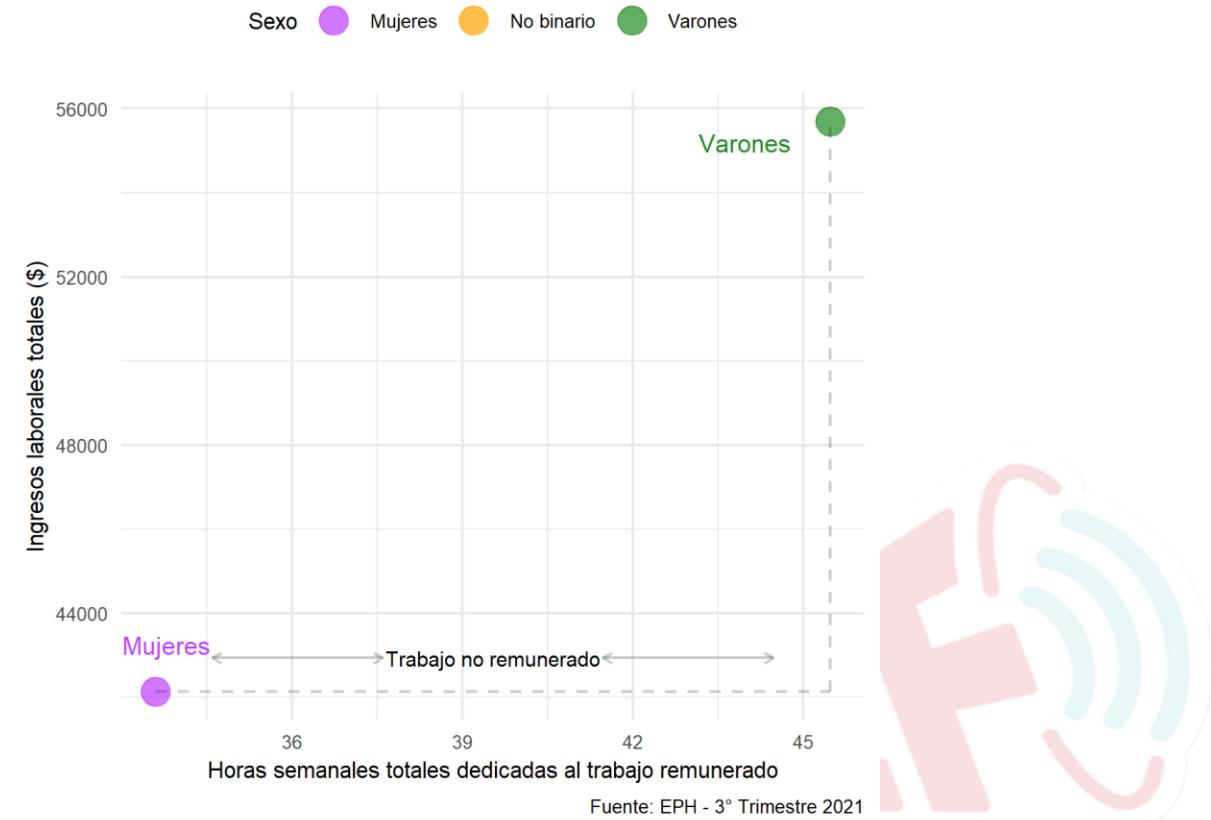
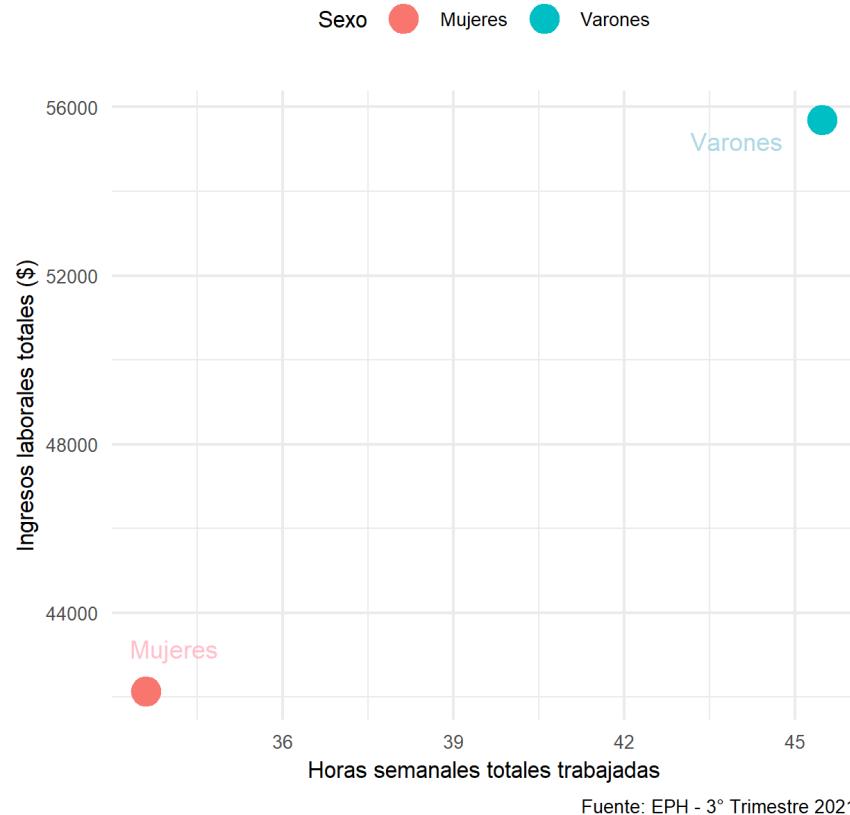
2. Análisis de los datos y comunicación de los resultados

- Discutir las convenciones y estereotipos
¿rosa y azul?
- Resaltar los espacios en blanco
- ¿Las emociones van en contra de la objetividad?



Etapas para **intervenir** sobre los sesgos y prejuicios de una investigación

2. Análisis de los datos y comunicación de los resultados



Notas

El término **minoría** no se utiliza para referir a segmentos de la sociedad con una cantidad reducida de miembros, sino a aquellos que son dominados y oprimidos por algún grupo que se encuentra en una situación dominante.

(Ej: varones vs. todas las demás identidades de género).



Referencias bibliográficas

- D'ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT press.
- Vazquez Brust, Antonio (2021), Ciencia de Datos para Gente Sociable
- R Core Development Team, (2000), Introducción a R. Notas sobre R: Un entorno de programación para Análisis de Datos y Gráficos.



Los contenidos han sido elaborados y revisados por el equipo de Ecofeminita y docentes del curso.

Coordinación:

Docentes:

Colaboración en desarrollo de contenidos:

www.ecofeminita.com



@EcoFeminita



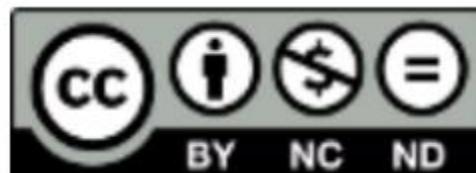
/ecofeminita



/ecofeminita



/channel/UCfykyqol8gBCdpohn4veXCA



Esta obra está bajo una Licencia Creative Commons Atribución – No Comercial – Sin Obra Derivada 4.0 Internacional. No se permite un uso comercial de la obra original ni la generación de obras derivadas.

