

SEM para medición económica y social 2024-I

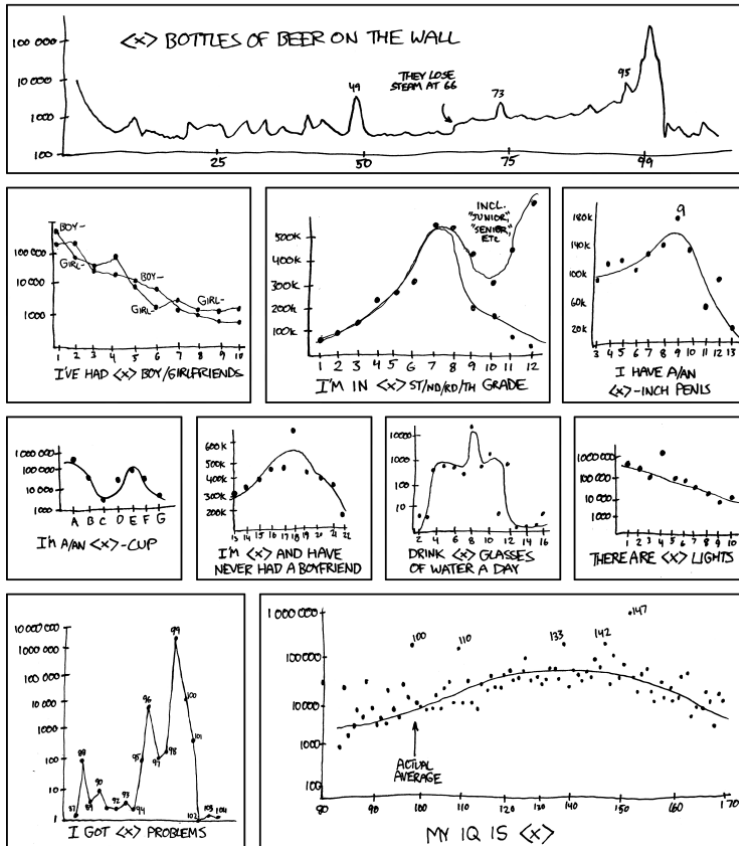
Dr. Héctor Nájera
PUED-UNAM



Los principios de la incertidumbre:

Stewart (2019) “The future is uncertain, but the science of uncertainty is the science of the future.”

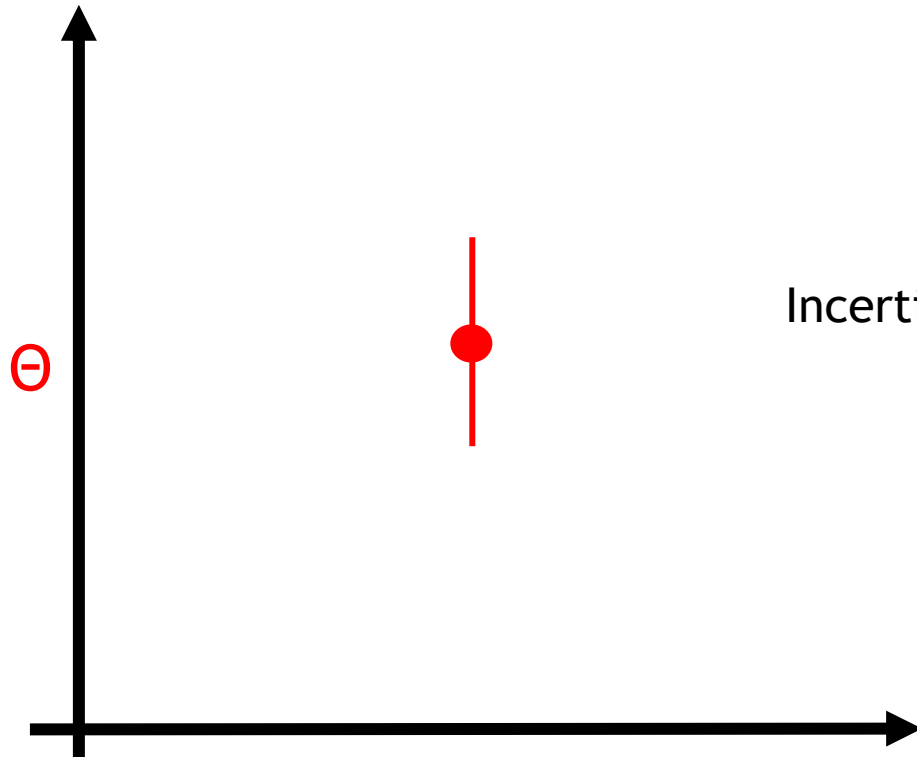
GOOGLE RESULTS FOR VARIOUS PHRASES:



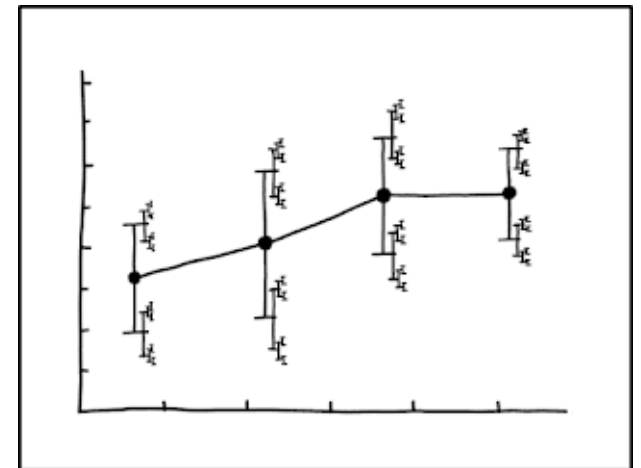
Aceptar la incertidumbre es una condición necesaria del buen análisis estadístico

Cuantificarla es el siguiente gran paso.

¿Incertidumbre?



Incertidumbre \neq Intervalo de confianza



I DON'T KNOW HOW TO PROPAGATE
ERROR CORRECTLY, SO I JUST PUT
ERROR BARS ON ALL MY ERROR BARS.

Curso sobre incertidumbre y medición

No es un curso convencional de estadística

No es un curso de recetas de cocina/estadística:

- Fórmula
- Receta y supuestos
- Interpretación



Principios del curso (Gelman et al., 2013):

“Statistics is said to be the science of defaults. One of our challenges is to defaultize things.”

“The full name of theoretical statistics is the theory of applied statistics.”

“Inference is not the inverse of a hypothesis test.”

“As you know from teaching introductory statistics, 30 is infinity.”

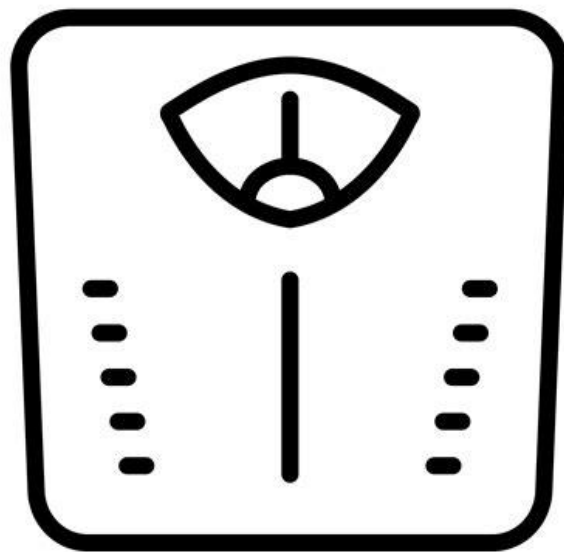
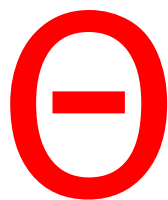
“It’s like the joint distribution is a movie, and all you care about is the star, like Robert Downey Jr. or whatever.”

“You can’t stand on the beach of the sea of uncertainty with the waves lapping at your ankles.

You have to jump into the sea and stick your head underwater and blow some bubbles.”

“Better to have analyzed and lost than never to have analyzed the data at all.”





Θ

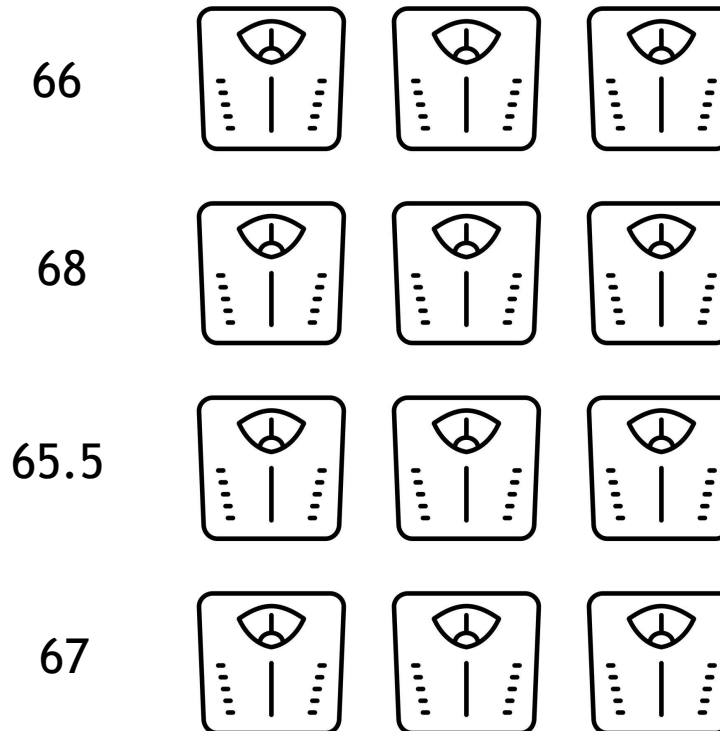


$\hat{\theta}$

¿Qué es el error estándar?

La desviación estándar del parámetro de interés

$\hat{\theta}$



$$\mu = 67.1$$

$$sd = .7$$

$$se = .7 / \sqrt{46} = .1$$

$$ci = 67.1 [66.9 - 67.3]$$

¿Usarían el CI para hablar de la incertidumbre de $\hat{\theta}$?

$$Y = \alpha + \beta_1 W + \beta_2 E + e$$

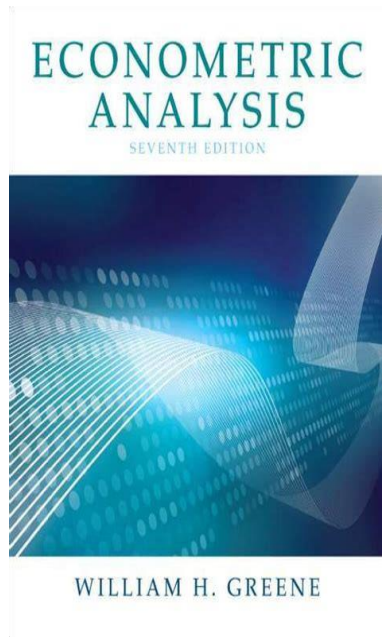


Noten que en realidad **W** viene de $\hat{\theta}$.

Noten que Y viene también de $\hat{\theta}_y$

Noten que E viene también de $\hat{\theta}_E$

$$Y = \alpha + \beta_1 W + \beta_2 E + e$$



Noten que en realidad **W** viene de $\hat{\theta}$.

Noten que Y viene también de $\hat{\theta}_y$

Noten que E viene también de $\hat{\theta}_E$

Θ









Instrumentos. Con distintas
fuentes de error.

$$\hat{\theta} = 50$$

$$\Theta - \hat{\theta} = 300$$

Números y rankings

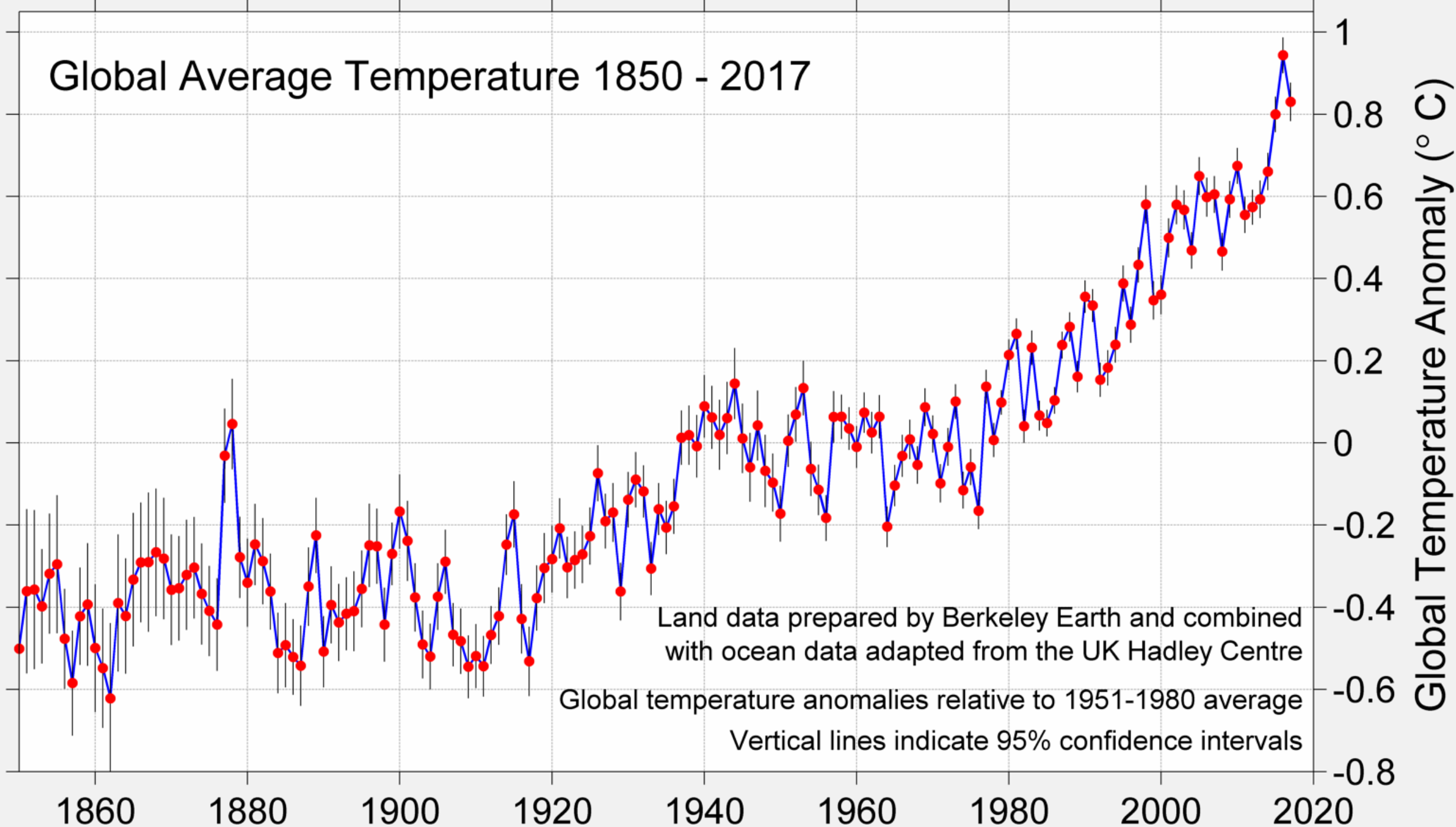


Rank	Bib	Name	NOC Code							Total
1	164	UCHIMURA Kohei	JPN	15.100	<u>15.066</u>	15.333	16.266	15.325	15.600	92.690
2	144	NGUYEN Marcel	GER	15.300	<u>13.666</u>	15.366	15.666	15.833	15.200	<u>91.031</u>
3	213	LEYVA Danell	USA	<u>15.366</u>	13.500	14.733	15.566	15.833	15.700	<u>90.698</u>
4	201	KUKSENKOV Mykola	UKR	<u>14.633</u>	14.600	15.200	15.533	15.400	15.066	<u>90.432</u>
5	193	BELYAVSKIY David	RUS	<u>14.466</u>	14.866	14.833	16.200	15.166	14.766	<u>90.297</u>
6	162	TANAKA Kazuhito	JPN	14.166	13.433	<u>15.200</u>	15.533	15.500	15.575	<u>89.407</u>
7	134	THOMAS Kristian	GBR	<u>15.566</u>	14.566	14.633	14.908	14.733	15.000	<u>89.406</u>
8	215	OROZCO John	USA	<u>15.433</u>	12.566	15.200	15.900	15.266	14.966	<u>89.331</u>
9	113	GONZALEZ Fabian	ESP	14.600	14.733	<u>13.966</u>	16.133	14.400	15.166	<u>88.998</u>
10	251	SASAKI JUNIOR Sergio	BRA	14.233	<u>14.366</u>	14.233	16.100	15.200	14.833	<u>88.965</u>
11	205	VERNIAIEV Oleg	UKR	14.533	13.966	<u>14.866</u>	16.233	15.033	14.300	<u>88.931</u>
12	222	SHATILOV Alexander	ISR	15.600	<u>14.266</u>	14.200	15.133	14.400	14.833	<u>88.432</u>
13	132	PURVIS Daniel	GBR	15.166	<u>14.266</u>	14.800	16.000	13.600	14.500	<u>88.332</u>
14	194	GARIBOV Emin	RUS	14.475	<u>14.233</u>	14.866	14.833	15.366	14.233	<u>88.006</u>
15	142	HAMBUCHEN Fabian	GER	<u>15.200</u>	13.266	14.800	14.766	15.400	14.333	<u>87.765</u>
16	125	TOMMASONE Cyril	FRA	13.500	15.333	<u>14.400</u>	15.358	15.000	14.066	<u>87.657</u>
17	245	CAPELLI Claudio	SUI	14.866	14.366	<u>14.166</u>	14.566	14.850	14.500	<u>87.314</u>
18	155	POZZO Enrico	ITA	14.700	13.900	14.000	<u>15.466</u>	14.533	14.433	<u>87.032</u>
19	227	JEFFERIS Joshua	AUS	14.066	13.533	14.800	<u>15.433</u>	14.900	14.133	<u>86.865</u>
20	174	KIM Soo Myun	KOR	12.266	13.700	14.200	<u>16.000</u>	14.641	14.966	<u>85.773</u>
21	254	VERBAEYS Jimmy	BEL	13.933	14.033	14.000	<u>15.266</u>	14.833	13.166	<u>85.231</u>
22	154	OTTAVI Paolo	ITA	12.466	14.033	15.016	<u>15.000</u>	14.100	14.033	<u>84.648</u>
23	112	GOMEZ FUERTES Javier	ESP	14.266	12.433	<u>14.800</u>	15.466	14.733	12.733	<u>84.431</u>
24	258	KULESZA Roman	POL	13.866	13.000	13.866	<u>14.400</u>	15.100	13.933	<u>84.165</u>

¿Qué significa estar 8 puntos abajo?

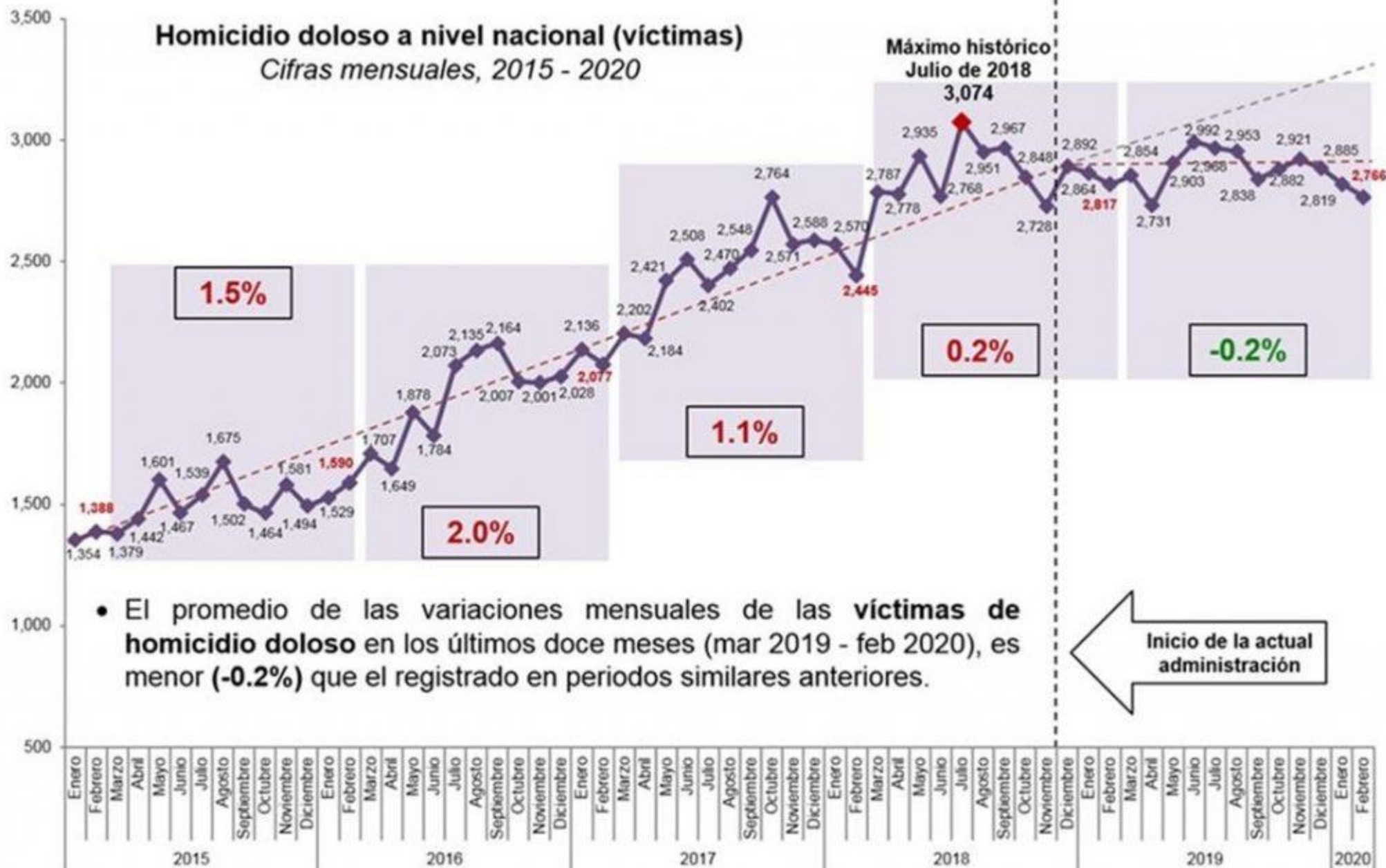
¿Qué nos permite hacer conclusiones sobre estos scores?

Global Average Temperature 1850 - 2017



Homicidio doloso a nivel nacional (víctimas)

Cifras mensuales, 2015 - 2020



- El promedio de las variaciones mensuales de las **víctimas de homicidio doloso** en los últimos doce meses (mar 2019 - feb 2020), es menor (**-0.2%**) que el registrado en periodos similares anteriores.

¿De verdad
2,766?

¿Números
absolutos?

¿Víctima de
homicidio doloso?

¿Víctima?

¿Homicidio
doloso?



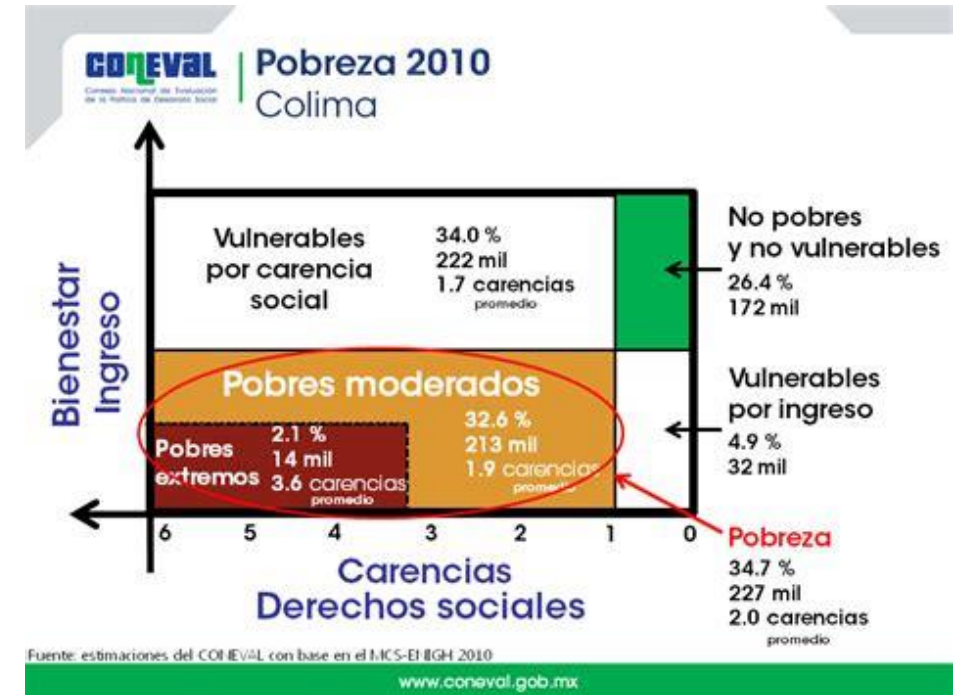
Números y rankings

No es muy distinto a lo que nos interesa hacer en distintas disciplinas

En qué términos $A > B$ y por tanto debe entrar a la UNAM?

Score A

Score B



La medición está en el corazón de la ciencia



- Nos permite concluir si cierto tratamiento es útil para abatir una enfermedad
- Podemos afirmar si el clima está cambiando y si es producto de la actividad humana
- Que la tecnología funcione y se comporte como esperamos que lo haga
- Diagnóstico de anticuerpos
- Flora y fauna en expansión o extinción
- Registrar señales de que las cosas están cambiando o permanecen en su estado actual
- Clasificar a las personas que aprovecharán de mejor manera los recursos universitarios
- Identificar si alguien tiene depresión aguda
- Si existen desigualdades sociales

CONCEPTOS/Abstracciones

Expectativas racionales

Precariedad laboral

Bienestar económico

Pobreza

Igualdad de oportunidades

Capital social

Desarrollo

Inversión social

Corrupción

Inflación

Igualdad de género

Capital cultural

Productividad

Pobreza multidimensional

Habilidad matemática

Bienestar social

Desempleo

Felicidad

Clase social

Informalidad

Calidad de un servicio

Inversión

Desarrollo



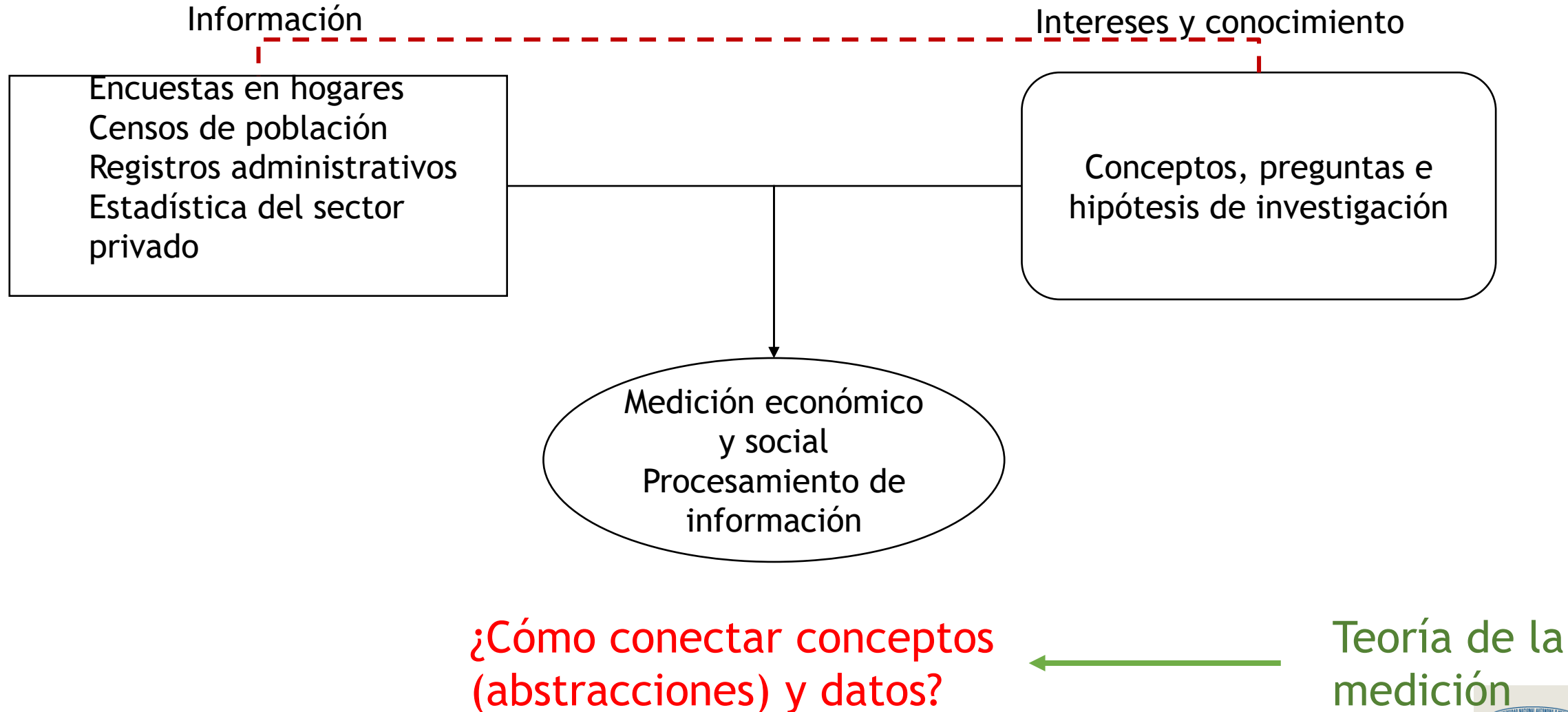
Preguntas

¿Cuál es la prevalencia de estos fenómenos?

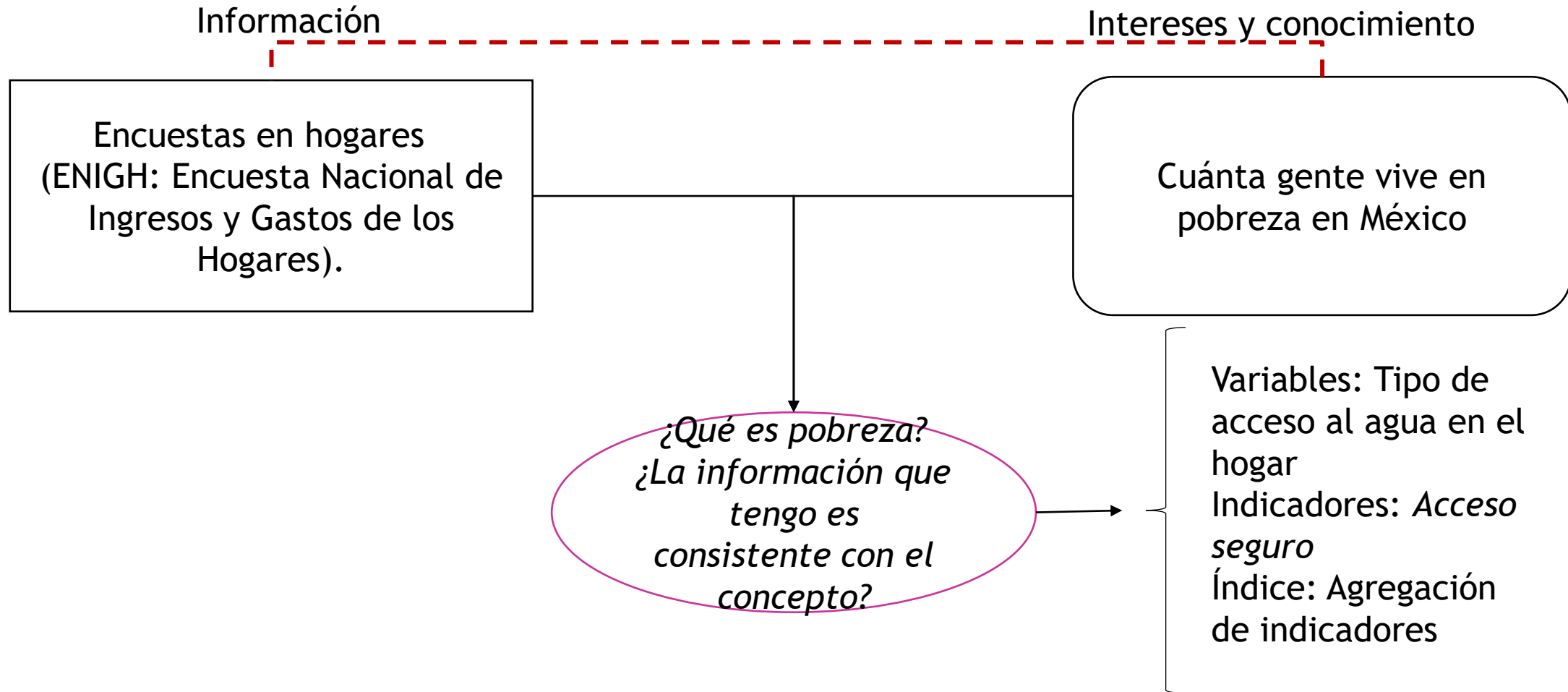
¿Cuál es su distribución poblacional, territorial o temporal?

¿Cómo se relacionan unos fenómenos con otros?

¿Por qué el curso?



Algunos ejemplos



¿Cómo sé que el índice resultante sirve para los propósitos de mi investigación?

Muchas ideas y muchos índices

¿Bajo qué criterios sabemos que podemos usarlos?

¿QUÉ ES EL ÍNDICE DE REZAGO SOCIAL?

Índice de Desarrollo Social de la Ciudad de México por manzana, 2021



Índice de Bienestar Económico Sostenible (IBES)
INAE V

Índice de progreso social

Economía

¿Cuál es la medida real de la informalidad?



CORRUPTION PERCEPTIONS INDEX

Índice de Desarrollo Humano (IDH): qué es y cómo se calcula

Social Welfare Index

WELFARE INDEX PMI 2021

Hacia un nuevo sistema de indicadores de bienestar



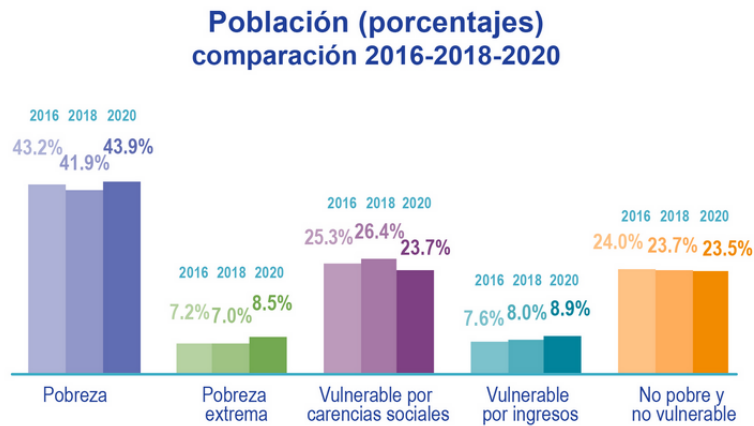
Mismo fenómeno diferentes resultados

MEDICIÓN DE POBREZA 2020

ESTADOS UNIDOS MEXICANOS



Bienestar económico	\$
	2016
	14.9%
Población con ingreso inferior a la línea de pobreza extrema por ingresos	2018
	14.0%
	2020
	17.2%
	2016
	50.8%
Población con ingreso inferior a la línea de pobreza por ingresos	2018
	49.9%
	2020
	52.8%



Carencias sociales



Human Development Reports

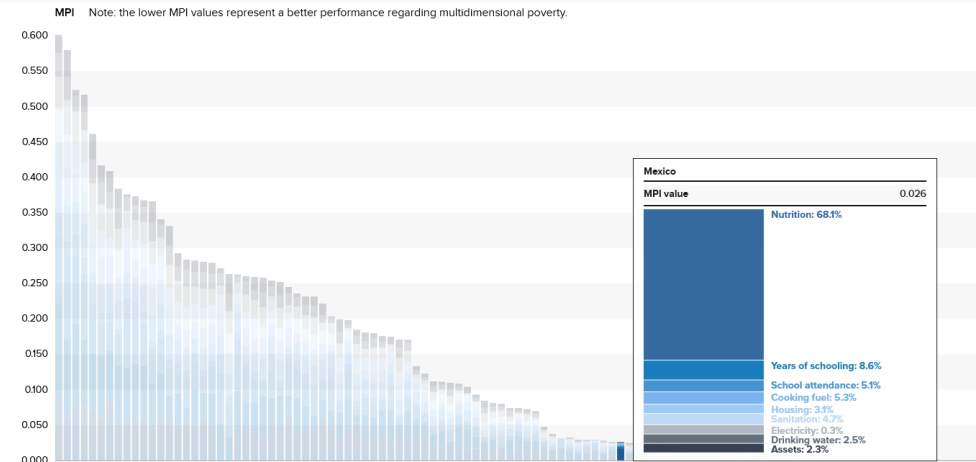
[HOME](#)

[REPORTS AND PUBLICATIONS](#)

[DATA CENTER](#)

[NEWS](#)

[ABOUT](#)



MPI Dimensions and Indicators



Estados Unidos Mexicanos

Distribución del porcentaje de la población por indicadores de pobreza, vulnerabilidad y no vulnerabilidad



* Para un mejor análisis de la información 2022, consultar las notas técnicas.

Estados Unidos Mexicanos

Indicadores de carencia social

	2016	2018	2020	2022*
Rezago educativo	18.5%	19.0%	19.2%	19.4%
Carencia por acceso a los servicios de salud	15.6%	16.2%	28.2%	39.1%
Carencia por acceso a la seguridad social	54.1%	53.5%	52.0%	50.2%
Carencia por calidad y espacios de la vivienda	12.0%	11.0%	9.3%	9.1%
Carencia por acceso a los servicios básicos en la vivienda	19.2%	19.6%	17.9%	17.8%
Carencia por acceso a la alimentación nutritiva y de calidad	21.9%	22.2%	22.5%	18.2%

* Para un mejor análisis de la información 2022, consultar las notas técnicas.

Estados Unidos Mexicanos

Líneas de Pobreza por Ingresos

Población con ingreso inferior a la línea de **pobreza** por ingresos

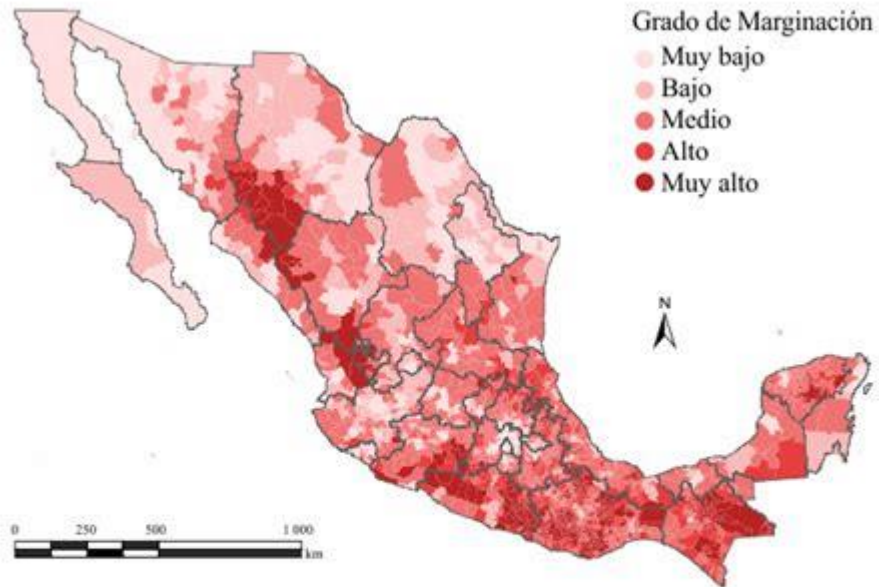
2016	2018	2020	2022*
50.8%	49.9%	52.8%	43.5%

Población con ingreso inferior a la línea de **pobreza extrema** por ingresos

2016	2018	2020	2022*
14.9%	14.0%	17.2%	12.1%

* Para un mejor análisis de la información 2022, consultar las notas técnicas.

Índice de Marginación



Fuente:
Elaboración propia con datos de:
Conapo. *Índice de Marginación por entidad federativa y municipio 2010*.
México, 2011. Base electrónica de datos. Fecha de consulta: mayo de 2014.

Índice de marginación (CONAPO, 2010)

¿Bajo qué criterios concluyo que los municipios en rojo tienen mayor marginación que los de color tenue?

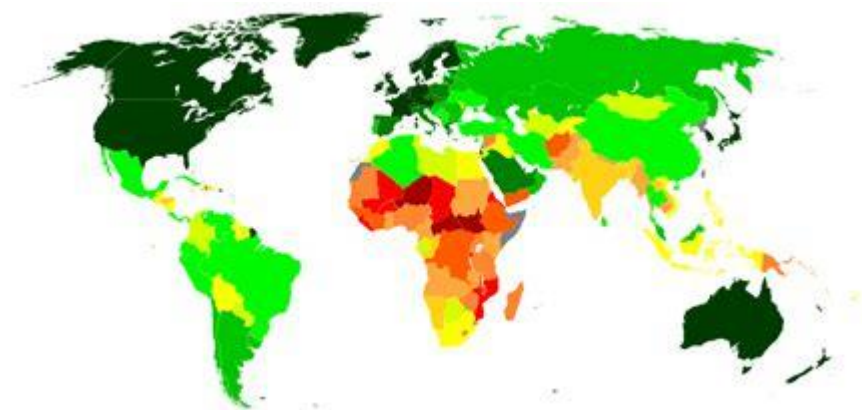
¿Qué me dice que esas cinco categorías son la mejor manera de representar los datos?

¿Hay algún indicador del índice que explica la variabilidad observada?

¿Cuál es la relación entre el concepto de marginación y el mapa?



Índice de desarrollo humano



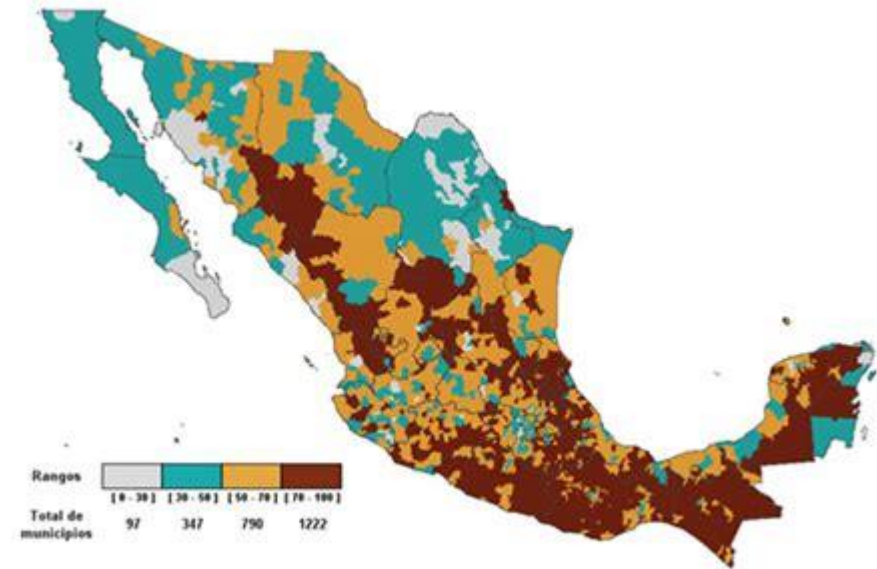
¿Bajo qué criterios puedo concluir que México, Chile, Argentina y Brasil tienen el mismo nivel de desarrollo humano?

¿El índice ordena lo que quiero que ordene (Mayor a menor desarrollo humano)?

Clasificación e inferencia



¿Qué me permite hacer inferencia sobre si la pobreza cayó, incrementó o se quedó igual?



¿Qué me permite hacer conclusiones sobre el patrón/distribución de la pobreza?

Múltiples variables y relaciones

Muchas veces nos interesa saber si un fenómeno se relaciona con otro (signo y magnitud de la relación)

¿Cómo puedo confiar en esos coeficientes?

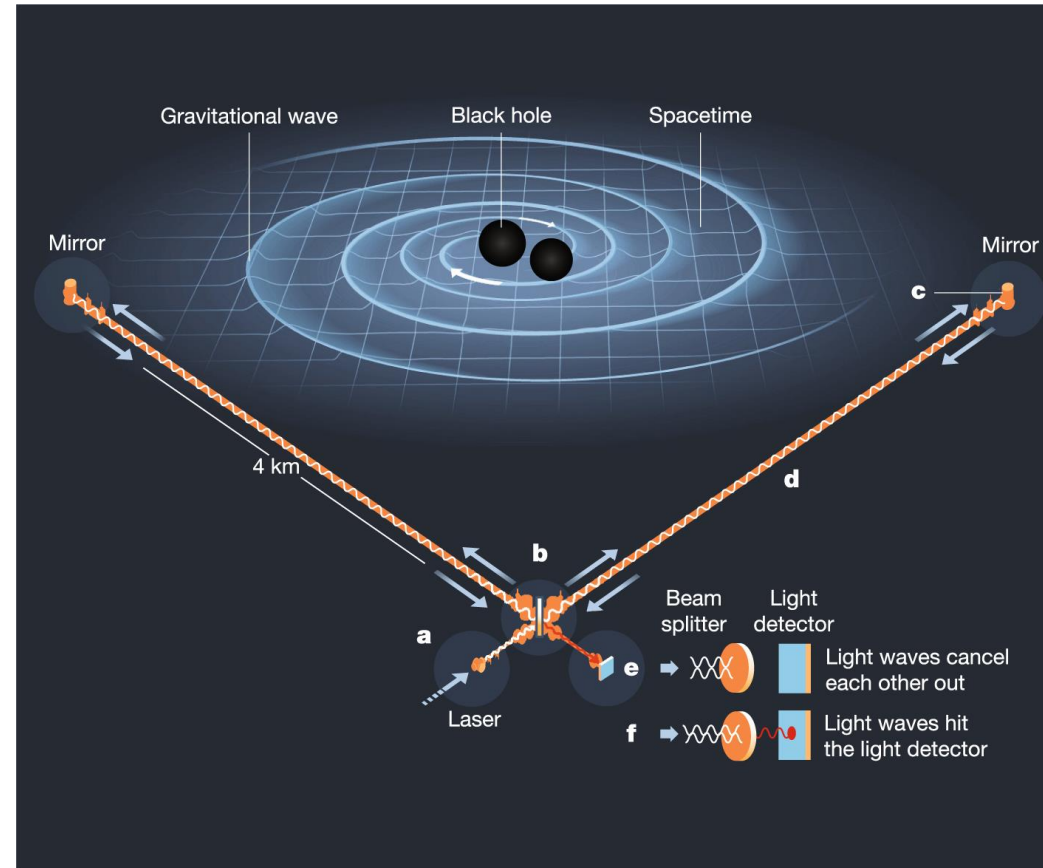
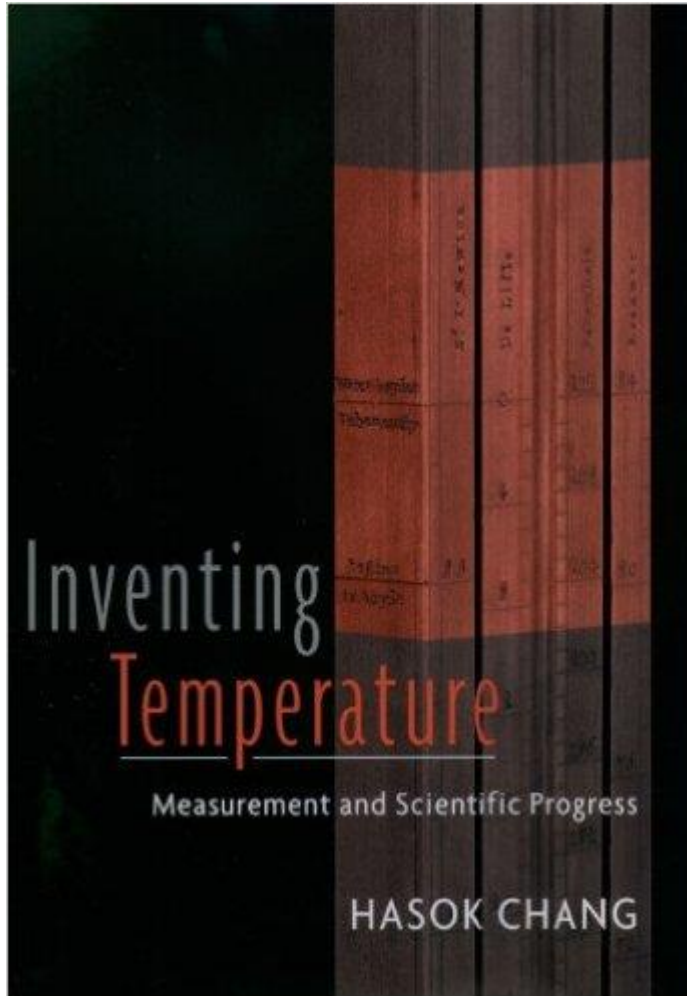
	Entry wage secondary education	10 years after secondary education	Entry wage tertiary education	10 years after tertiary education	Perceived returns to schooling
Female	-0.252** (0.083)	-0.347*** (0.089)	-0.252*** (0.076)	-0.395*** (0.080)	0.136 (0.096)
Age	0.059 (0.056)	0.119* (0.059)	0.047 (0.055)	0.018 (0.052)	-0.041 (0.066)
Immigrant	0.086 (0.096)	0.064 (0.088)	0.101 (0.071)	0.086 (0.083)	0.076 (0.134)
Father's education low	0.266 (0.179)	0.377** (0.142)	0.441** (0.162)	0.211 (0.147)	-0.231 (0.329)
Father's education medium	0.189** (0.072)	0.164* (0.068)	-0.017 (0.071)	-0.044 (0.072)	-0.177 (0.118)
Mother's education low	-0.032 (0.122)	-0.159 (0.149)	-0.132 (0.133)	-0.189 (0.167)	0.240 (0.279)
Mother's education medium	0.051 (0.067)	0.113 (0.067)	0.121 (0.067)	0.008 (0.070)	0.058 (0.105)
Not proceeding to University	-0.065 (0.097)	-0.048 (0.103)	-0.198** (0.076)	-0.178* (0.078)	-0.542*** (0.115)
Next-to-last year students	0.115 (0.096)	0.106 (0.103)	0.090 (0.105)	-0.009 (0.101)	-0.189 (0.108)
Probability of completion	-0.007** (0.002)	-0.007** (0.002)	-0.001 (0.002)	-0.001 (0.002)	0.006 (0.004)
Constant	7.025*** (0.440)	7.214*** (0.460)	7.296*** (0.432)	8.112*** (0.431)	3.222*** (0.570)
R ²	0.055	0.078	0.042	0.061	0.070

Muchos datos no implica mejor medición



Depende qué hago con los datos para vincularlos al espacio de teorías y conceptos.

Pero es ciencias sociales...



SEM y medición socioeconómica

Andrew Gelman: ¿Qué tema en estadística está ausente en la mayoría de los textos?

Variación, comparación o medición

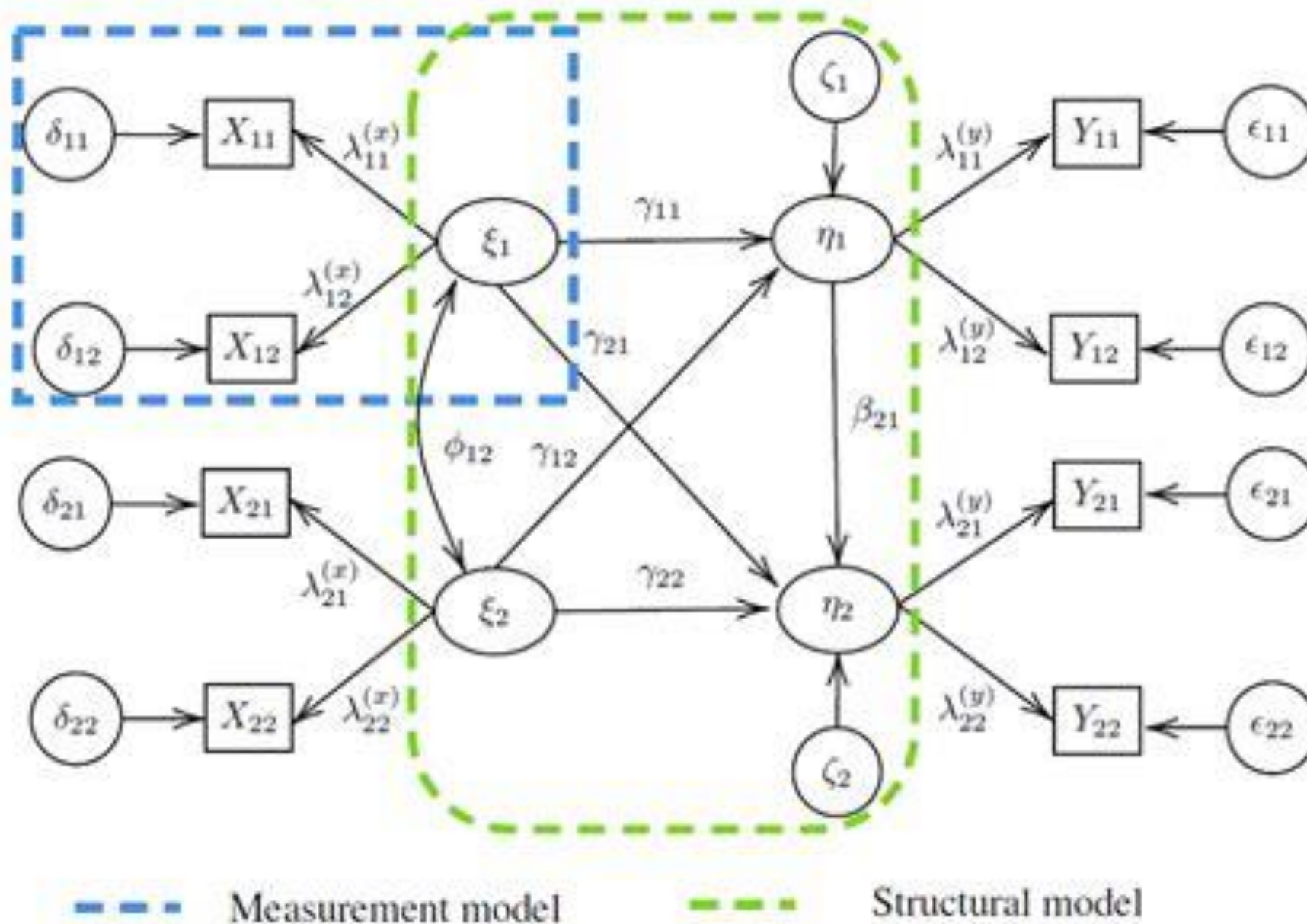
La mayoría de los textos en estadística cubren análisis de datos y muestreo, pero raramente lo integran con medición

SEM es, probablemente, el mejor marco analítico para estimar el error de medición socio-económica.



SEM

- Structural Equation Modelling
- Más de 100 años de constante evolución que han dado origen a un marco unificado con modelos no excluyentes sino complementarios:
 - Análisis factorial
 - Modelos de senderos
 - Modelos de ecuaciones múltiples
 - Modelos anidados o multinivel



SEM

Objetivo

Desarrollar las capacidades críticas y analíticas de los estudiantes para la producción y escrutinio de **índices** sociales e **indicadores** económicos.



Expectativas del curso

Al final del curso la es que los alumnos sean capaces de:

Entender por qué es importante trabajar con medidas falsables en ciencias sociales

Identificar la diferencia entre **distintas tradiciones** en medición

Comprender por qué la medición basada en **modelos (teórico-estadísticos)** es la mejor manera de abordar el problema en cuestión

Distinguir entre un método de agregación y una metodología de escrutinio empírico

Comprender las implicaciones de una mala medición en estadística inferencial

Apreciar la relevancia de la teoría de la medida para producir y examinar índices sociales

Comprender los vínculos entre la teoría de la medida, variables latentes y ecuaciones estructurales

Entender por qué los principios de confiabilidad y validez son una necesidad necesaria para una calidad mínima de medición

Implementar análisis de *ecuaciones estructurales* de confiabilidad y validez usando: **R-software**

Interpretar los resultados de los análisis de una forma crítica

Identificar los usos apropiados e inapropiados de ecuaciones estructurales



Plan de trabajo

- Observación, medición y error
 - Qué es medir y qué es error de medición
- Principios de medición
 - Error de medición, su historia y estimación
- Principios de medición: **Confiabilidad**
 - Conceptos y métodos de cálculo
 - Clásica y ecuaciones estructurales
- Principios de medición: **Validez**
 - Conceptos y métodos de cálculo
 - Clásica y ecuaciones estructurales
- Ecuaciones estructurales para el análisis de datos
 - Modelos de relaciones múltiples (Path analysis)
 - Modelos de clasificación (Latent Class)
 - Modelos multinivel

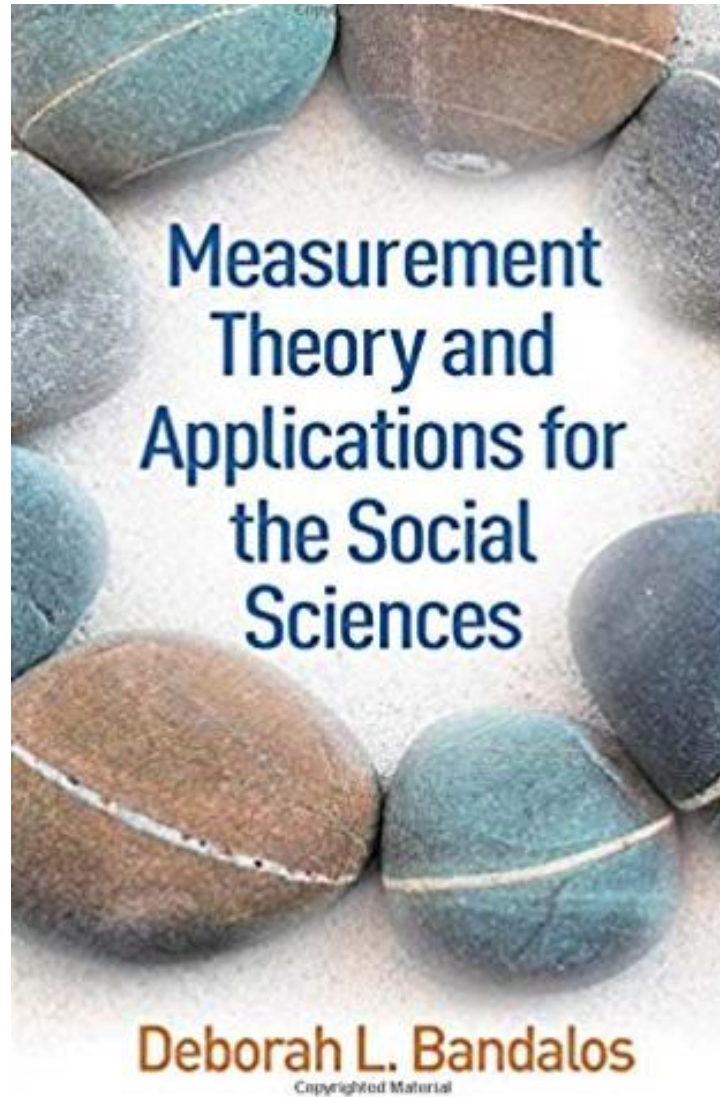


Características de las sesiones

- Las sesiones combinan discusión, teoría y aplicación con el programa R.
- Los docentes impartirán cada sesión (prepararán un archivo.ppt) y se dedicará siempre un espacio para discusión, ejercicios en grupo y/o implementación de análisis usando el programa R.



Bibliografía básica



Evaluación

Dos ejercicios:

1. Ejercicio sobre confiabilidad con SEM (50%)
2. Ejercicio sobre validez con SEM (50%)



Materiales

- Github:
- <https://github.com/hectornajera83/SEM2024I>

Próxima clase

Lecturas para la siguiente sesión

- Esencial
 - Fry, Hanna 2021. *What data can't do*. The New Yorker: <https://www.newyorker.com/magazine/2021/03/29/what-data-cant-do>
 - Loken, E., & Gelman, A. (2017). Measurement error and the replication crisis. *Science*, 355(6325), 584-585.



CONTACTO

Dr. Héctor Nájera
Investigador (SNI-I)

Programa Universitario de Estudios del Desarrollo (PUED)
Universidad Nacional Autónoma de México (UNAM)

Antigua Unidad de Posgrado (costado sur de la Torre II de Humanidades),
planta baja.

Campus Central, Ciudad Universitaria, Ciudad de México, México.

Tel. (+52) 55 5623 0222, Ext. 82613 y 82616

Tel. (+52) 55 5622 0889

Email: hecatalan@hotmail.com

*¡Bienvenidos
estudiantes!*



PONTE
PUMA



PONTE & CUBREBOCAS PUED
UNAM