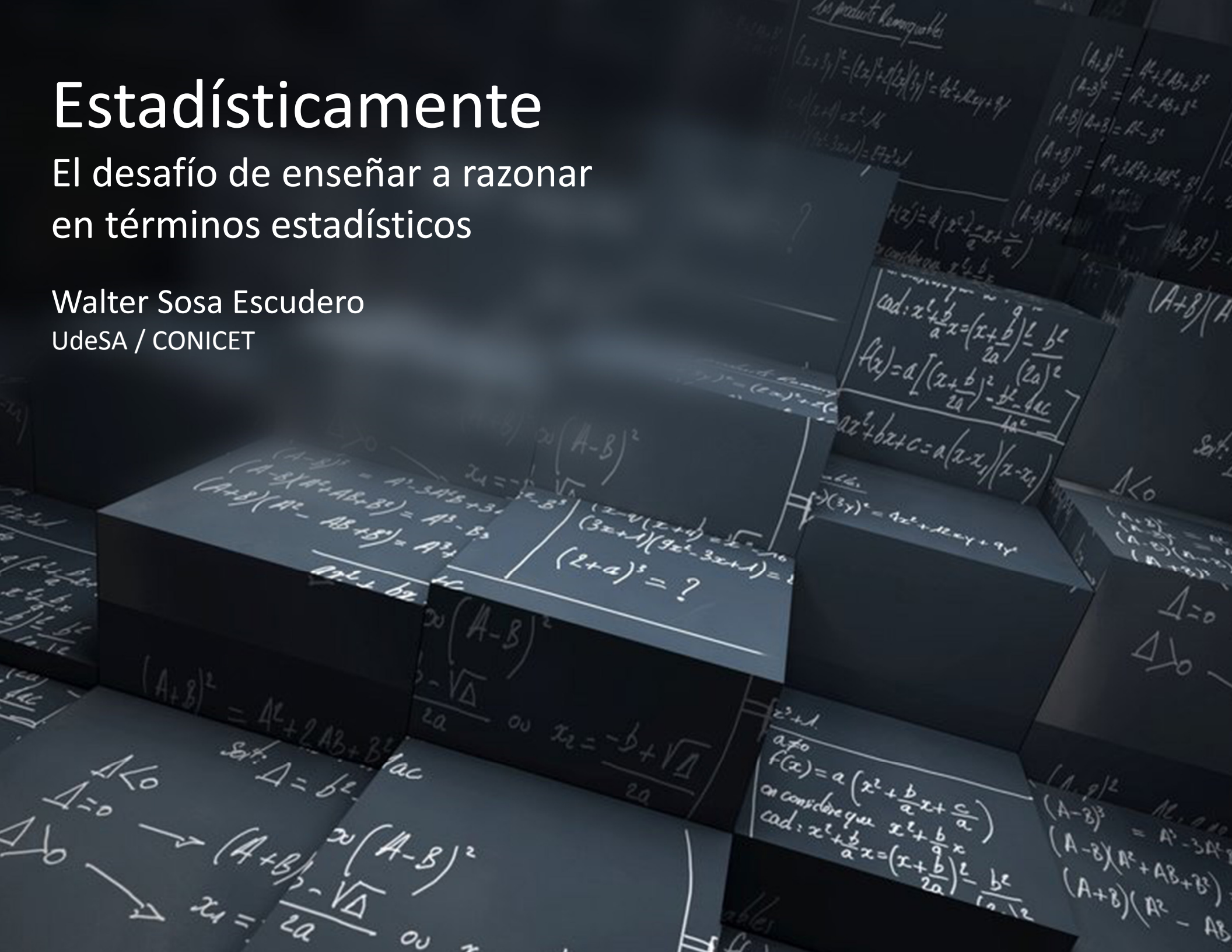
Estadísticamente

El desafío de enseñar a razonar en términos estadísticos

Walter Sosa Escudero
UdeSA / CONICET



Sexo: MUJER 19 julio 2012

Metabolismo Lipídico

Resultado	Unidad
221	mg/dl
63	mg/dl
108	mg/dl
100	mg/dl

Proteinas

Resultado	Unidad

Información nutrimental

Porción 1 manzana grande (242g / 8 oz.)

Cantidad por porción

Calorías 130

Grasas totales 0g

Grasas saturadas 0g

Grasas trans 0g

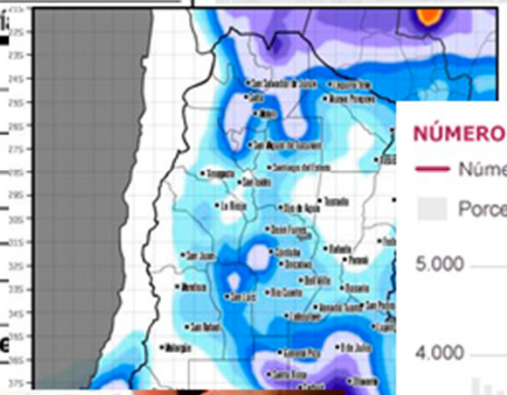
Colesterol 0mg

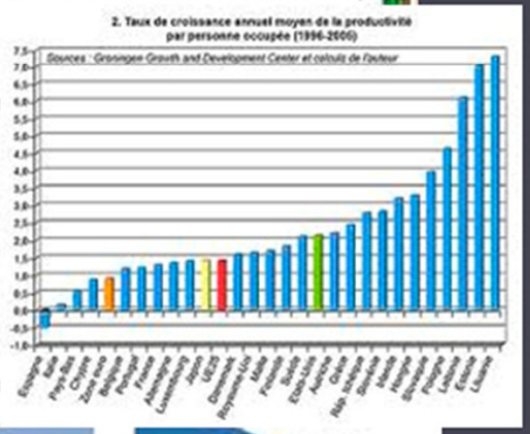
Sodio 0mg

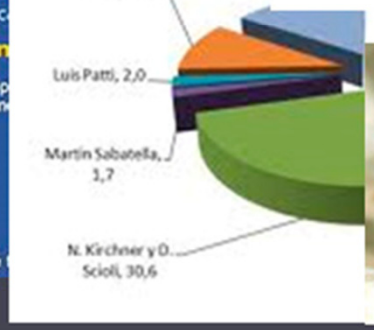
Potasio 260mg

Carbohidratos totale

Fibra dietética 5g

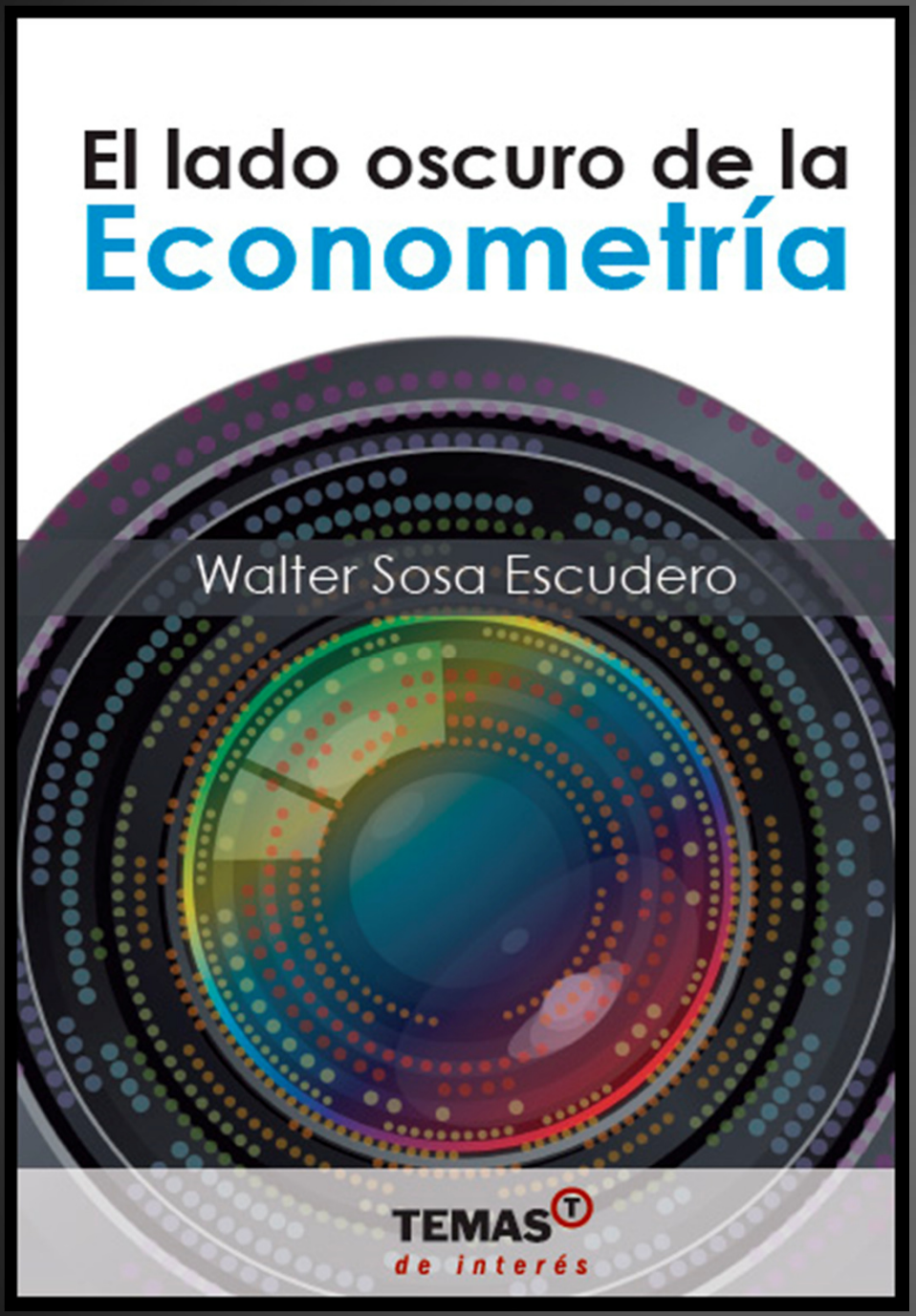


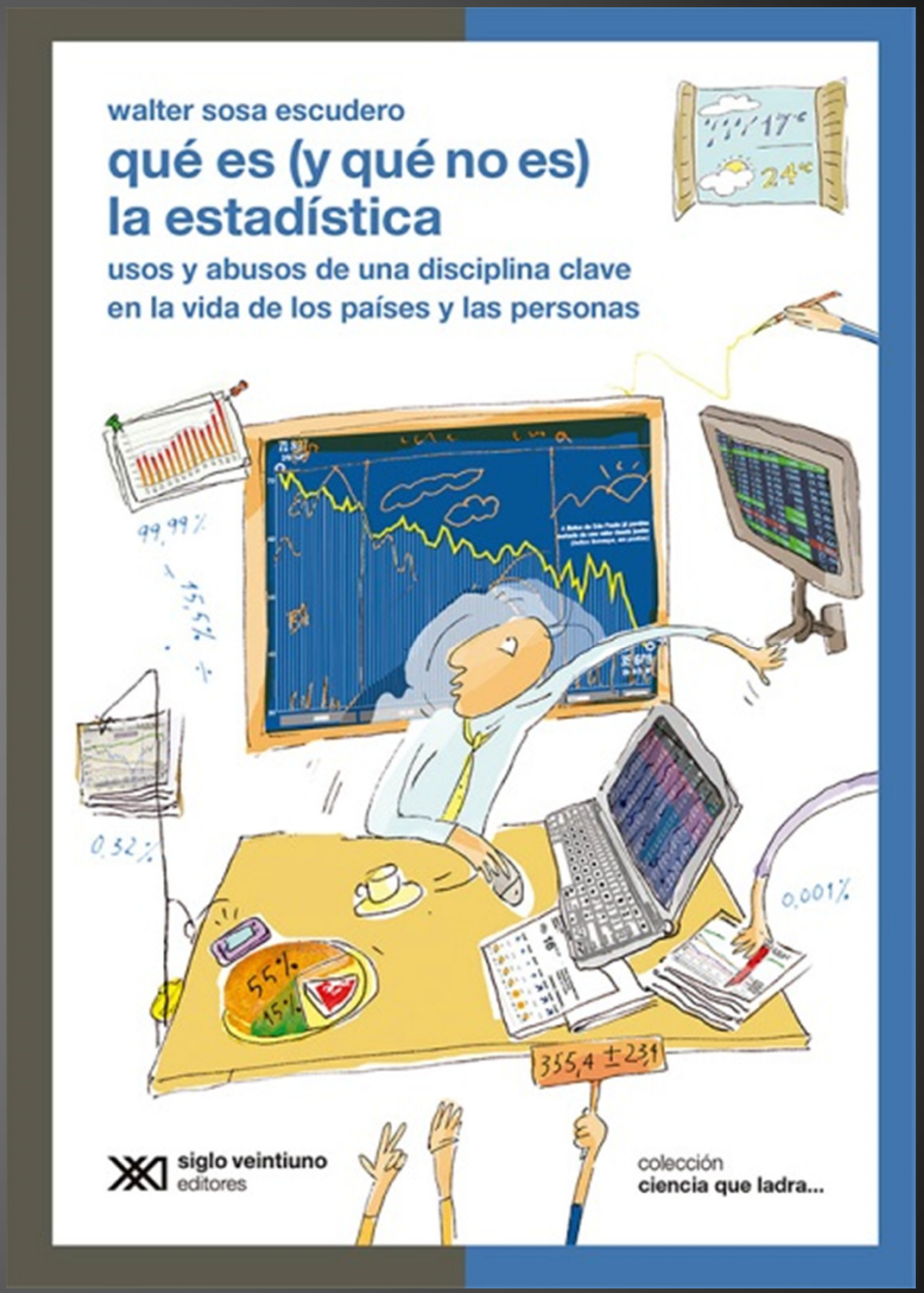




El desafío de enseñar a razonar en términos estadísticos

- Técnicas, algoritmos, métodos.
- La estadística como *instinto*.
- Desafío: la conveniencia de razonar en términos estadísticos.
- Estimar y predecir, por sobre conocer.
- Ejemplo: desempleo. Censos, población infinita, superpoblaciones.

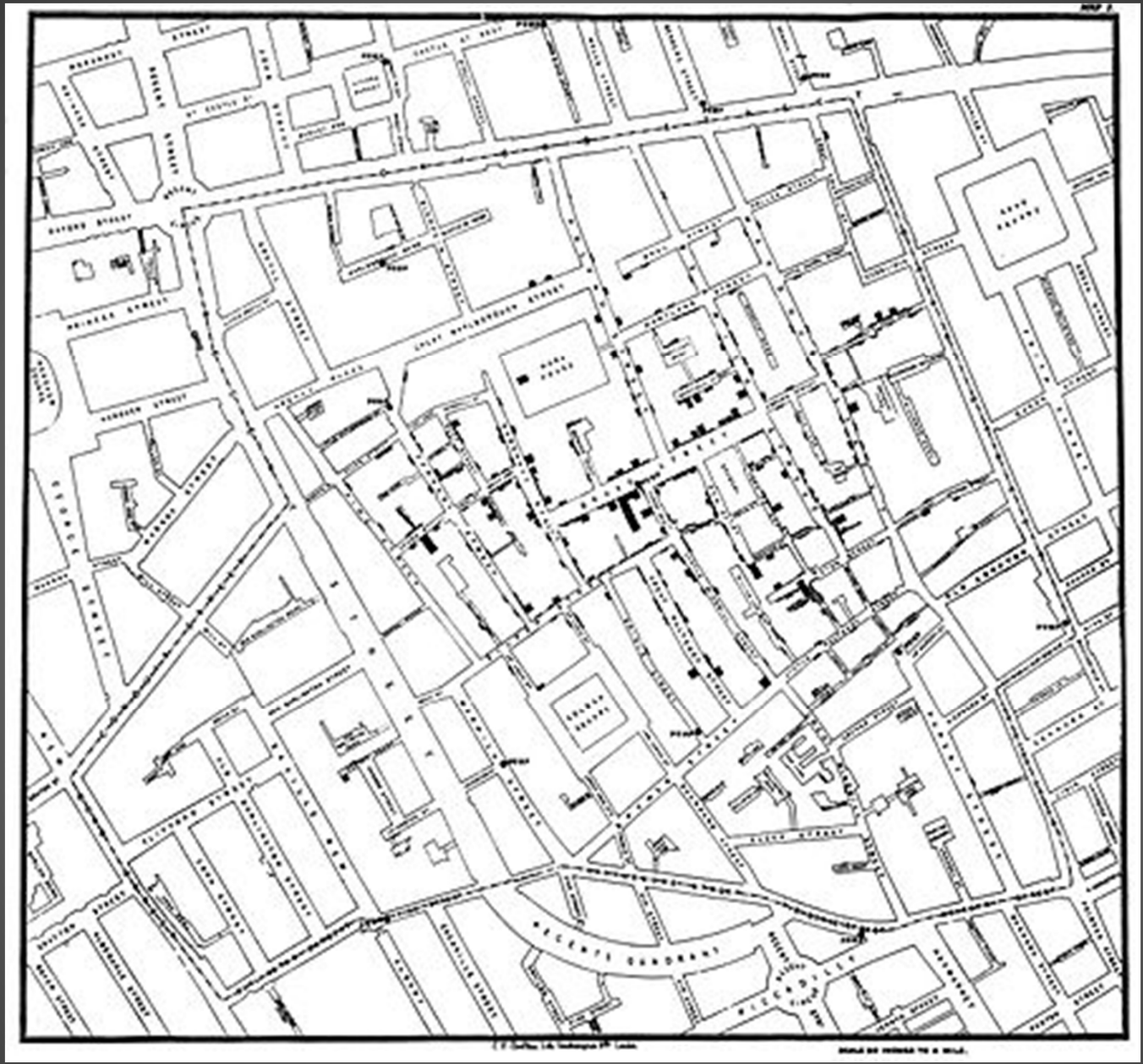

El lado oscuro de la
Econometría
Walter Sosa Escudero
TEMAS
de interés


walter sosa escudero
qué es (y qué no es)
la estadística
usos y abusos de una disciplina clave
en la vida de los países y las personas
99,99%
45,5%
0,32%
0,001%
55%
355,4 ±23,1
siglo veintiuno
editores
colección
ciencia que ladra...

Tres Casos

La estadística en los tiempos del cólera

- Año 1854, Soho, Londres.
- 31 de agosto: John Snow reporta un brote de colera en el Soho, Londres.
- 3 de septiembre: 127 muertos
- 6 de septiembre: 500 muertos
- 7 de septiembre: milagro. Gran logro de la epidemiologia, la política sanitaria… y la estadística.
- Joya de la comunicación visual (Ed Tufte)



El mapa original de Snow

El google map de Snow

Razonamientos

- Argumento estadístico vs. causal
- Argumento visual

- Referencias

  - Shiode N, Shiode S, Rod-Thatcher E, Rana S, Vinten-Johansen P. The mortality rates and the space-time patterns of John Snow's cholera epidemic map. *International Journal of Health Geographics*. 2015;14:21. doi:10.1186/s12942-015-0011-y.
  - Koch, T., 2004, The Map as Intent: Variations on the Theme of John Snow. C*artographica: The International Journal for Geographic Information and Geovisualization,* 39:4, 1-14
  - Analisis en R: http://www.r-bloggers.com/interactive-maps-for-john-snows-cholera-data/
  - http://www.ph.ucla.edu/epi/snow.html

Sexo, drogas y estadísticas

Exámen: marque con “x” la respuesta adecuada.
En el último año ¿Ha consumido Ud. alguna droga ilegal?

Si [ ]

No [ ]

- Pregunta delicada
- Proclive a no respuestas o respuestas mentirosas
- Drogas, racismo, evasión, delitos.

¡La estadística al rescate!

Instrucciones

A cada persona le damos un dado y una moneda.

1. Tire el dado.
2. Si sale 1 o 2, tire la moneda. Si sale cara, responda “si”. Si sale ceca, responda “no”.
3. Si sale 3, 4, 5 o 6, responda la verdad.

El método

- W = proporción de personas que responde “si” a nuestro experimento (mienta o no).
- P = proporción de consumidores en el grupo

$$P \approx W * 1.5 - 0.25$$

Ejemplo: si 30% responden que si, $P \approx 0.3*1.5-0.25 = 0.2$ (20% de consumidores).

Encuesta = 2/6 Moneda + 4/6 Verdad

Despejamos "Verdad":

Verdad = 6 / 4 Encuesta – 2 / 4 Moneda

Moneda = 1 / 2

Verdad = 1.5 Encuesta – 0.5 x 0.5

Verdad = 1.5 Encuesta – 0.25

Computar para creer…

```
Experimento sexo, drogas, estadistica

#valor verdadero = 1/3
x<-c(rep(1,30), rep(0,60))
dado<-sample(c(1:6), 90, rep=T)
coin<- sample(c(1,0), 90, rep=T)
rp<-x
rp[dado<3]<-coin[dado<3]
ag<-mean(rp)
1.5*ag-0.25

#Experimento de Montecarlo
expe<-rep(0,1000)
for(i in 1:1000){
 dado<-sample(c(1:6), 90, rep=T)
 coin<- sample(c(1,0), 90, rep=T)
 rp<-x
 rp[dado<3]<-coin[dado<3]
 ag<-mean(rp)
 expe[i]<-1.5*ag-0.25
 }
mean(expe)
```

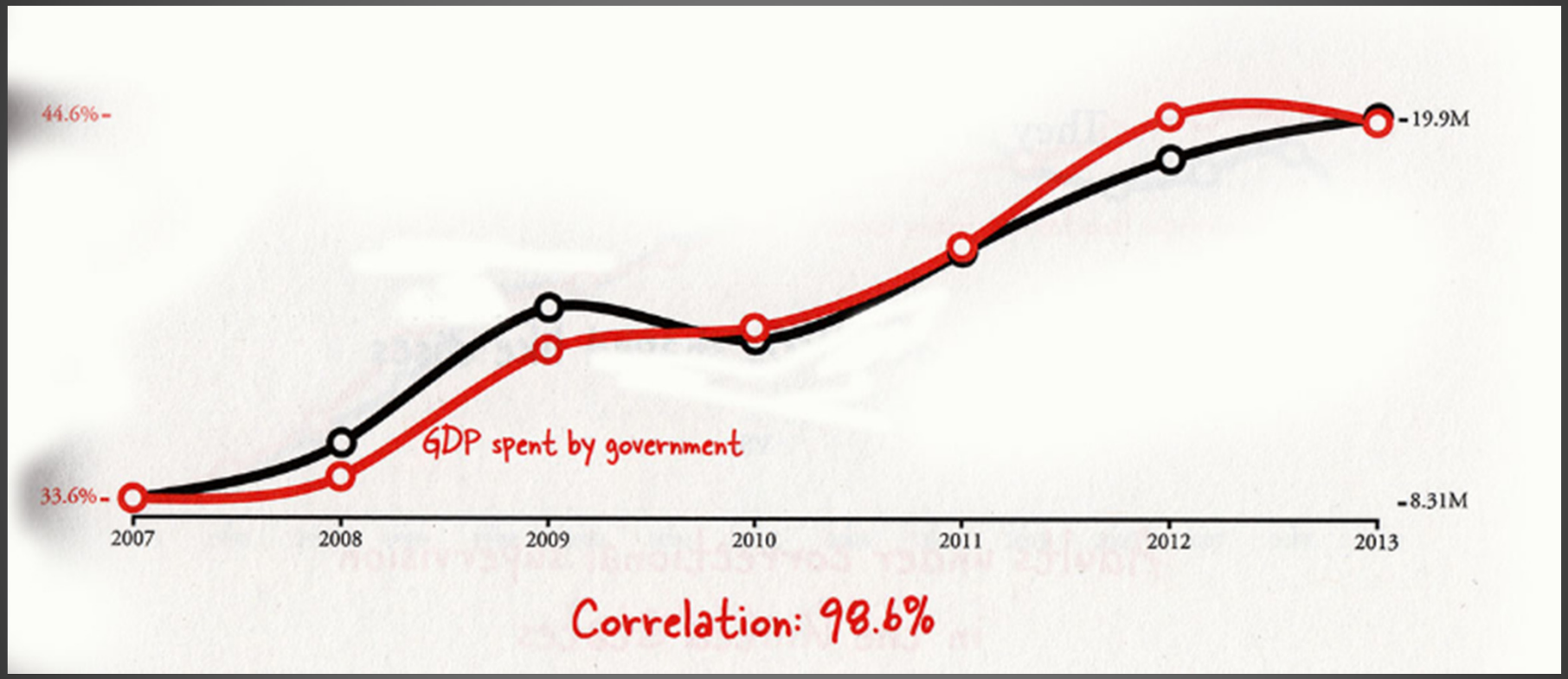
Razonamientos

- Cambiar el problema.
- Cuestiones complejas, estadísticas complejas.
- Tirar datos a la basura, métodos a la basura.

- Referencias

  - Bolstad, W., L. Hunt y J. McWhirter , 2001, "Sex, Drugs, and Rock and Roll survey in a First-Year Service Course in Statistics", The American Statistician, 55(2): 145-149
  - Alaimo, V. y Sosa Escudero, W., 2000,, La Economía Oculta en la Argentina: Evidencia Basada en Encuestas de Gasto, 2000, en FIEL, La Economia Oculta en la Argentina.

Euglossa bazinga

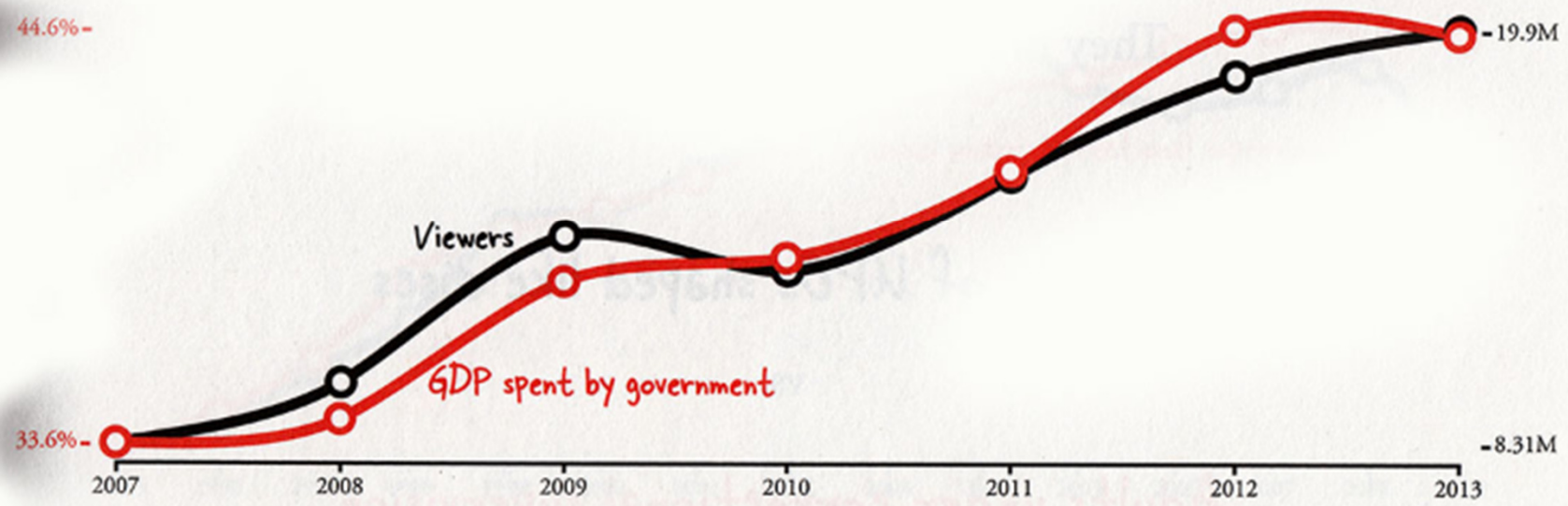

44.6%
33.6%
19.9M
8.31M
GDP spent by government
2007
2008
2009
2010
2011
2012
2013
Correlation: 98.6%

Bazinga, Argentina!

The Big Bang Theory viewership

vs.

Percentage of Argentina's GDP spent by the government



Correlation: 98.6%

The *Euglossa bazinga* is a species of bee in Argentina. That's not spurious: The bee was named after Sheldon's favorite word.

Sources:
Wikipedia: "The Big Bang Theory"
Smithsonian Magazine, "A Brand New Bee Was Just Named After Sheldon From *The Big Bang Theory*"
International Monetary Fund, World Economic Outlook Database

x=Nicolas cage films

y=swimming pool drownings

x=Divorces in Maine

y=Margarine consumption

Spurious correlations

CORRELATION DOES NOT EQUAL CAUSATION

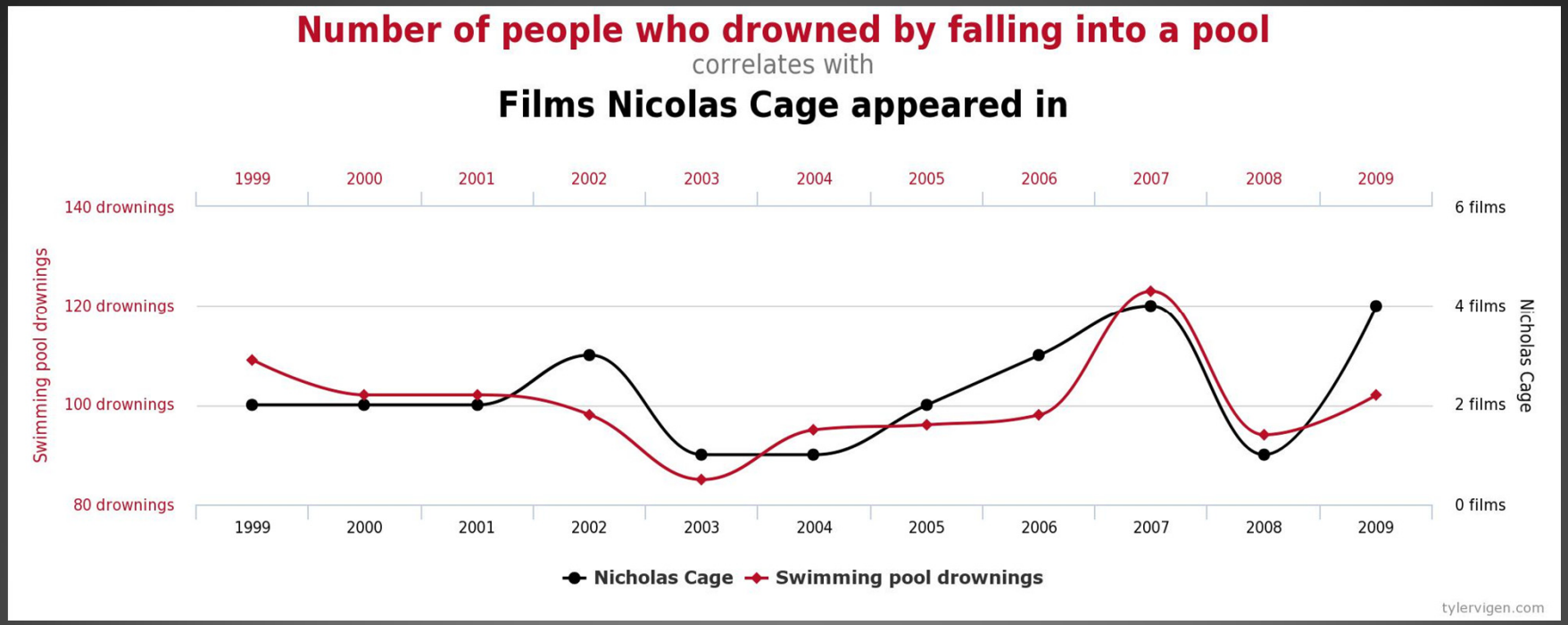
x=cheese eating

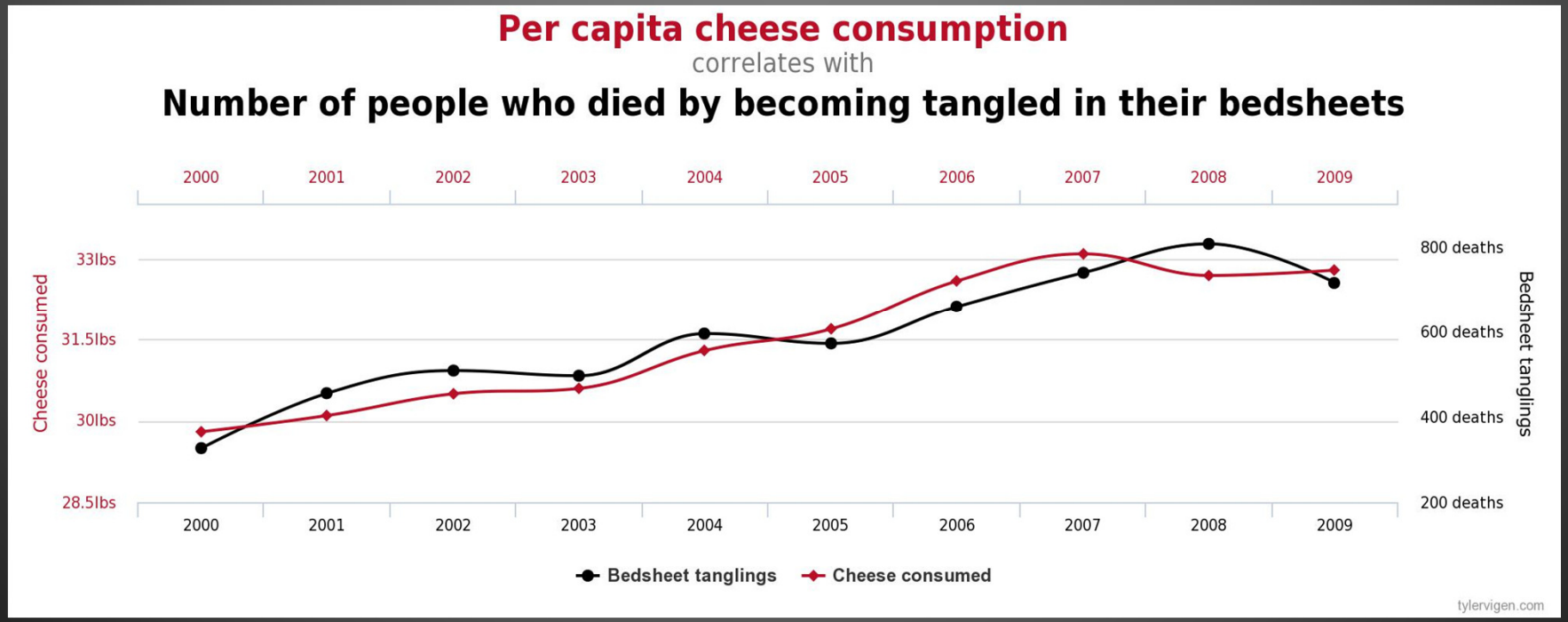
y=Fatal Linen Tangles

TYLER VIGEN

x=shark attacks

y=tornadoes


Number of people who drowned by falling into a pool
correlates with
Films Nicolas Cage appeared in
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
140 drownings
120 drownings
100 drownings
80 drownings
Swimming pool drownings
6 films
4 films
2 films
0 films
Nicholas Cage
Nicholas Cage
Swimming pool drownings
tylervigen.com


Per capita cheese consumption
correlates with
Number of people who died by becoming tangled in their bedsheets
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
33lbs
31.5lbs
30lbs
28.5lbs
Cheese consumed
800 deaths
600 deaths
400 deaths
200 deaths
Bedsheet tanglings
Bedsheet tanglings
Cheese consumed
tylervigen.com

Razonamientos

- El mensaje: ¿Casualidad (Neyman-Pearson), causalidad, correlación?
- El mensajero: ¿Correlacion espuria?
- ¿Tyler Vigen?

- Referencias

  - Granger, C. y Newbold, P., 1974, Spurious regressions in econometrics, Journal of Econometrics, 2, 111-120.
  - Phillips, P., 1986, Understanding spurious regressions in econometrics, Journal of Econometrics, 33, 311-340.

El desafio de pensar en términos estadísticos

Reflexiones

- Box: “All models are wrong, but some are useful”
- Lo bueno, lo malo, *útil* y lo *inutil*
- La matematica y la contabilidad no toleran errores, la estadística vive de los errores.
- La sofisticación de los razonamientos estadísticos y su utilidad.
- Divulgación científica.

Sugerencias

- Nate Silver. Five Thirty Eight: fivethirtyeight.com
- Andrew Gelman (blog): Statistical Modeling, Causal Inference and Social Science.
- Gelman y Nolan: Teaching Statistics: A Bag of Tricks.
- Scheaffer et al.: Activity based statistics.
- Bad Science (www.badscience.net).
- Econometrics beat (David Giles).

Econometria Avanzada
https://www.facebook.com/groups/econometriaudesa/
Walter
Inicio
Walter Sosa
Editar perfil
FAVORITOS
Noticias
Mensajes 20+
Eventos 2
Guardado
Big Data
GRUPOS
Econometria Ava...
Polémica Online 1
Comunidad La Aldea
Comunidad Argent...
Centro de Exalumn... 2
Econometría I
Yo conocí el verda...
Nuevos grupos
Crear grupo
APLICACIONES
Juegos 20+
Un día como hoy
Econometria Avanzada
(6b) $E(Y_{2i}|X_{2i}, Y_{2i} \geq 0) = X_{2i}\beta_2 + \frac{\sigma_{22}}{(\sigma_{22})^{1/2}} \lambda (Z_i)$
!!
Grupo público
Eres miembro
Compartir
Notificaciones
Debate
Miembros
Eventos
Fotos
Archivos
Buscar en este grupo
Publicar
Foto/video
Preguntar
Archivo
Escribe algo...
MIEMBROS 11 360 miembros (141 nuevos)
+ Agregar personas al grupo
Invitar por correo electrónico
PUBLICACIÓN MARCADA
Walter Sosa
14 de octubre de 2014
Bienvenida y algunas reglas: Estimados, sorprendido gratamente por la masividad de este grupo, les doy la bienvenida a los recien intergrados.
DESCRIPCIÓN
Editar
Hola. Este es nuestro espacio en Facebook....
Ver más

"En aquel Imperio, el Arte de la Cartografía logró tal Perfección que el Mapa de una sola Provincia ocupaba toda una Ciudad, y el Mapa del Imperio, toda una Provincia. Con el tiempo, estos Mapas Desmesurados no satisficieron y los Colegios de Cartógrafos levantaron un Mapa del Imperio, que tenía el Tamaño del Imperio y coincidía puntualmente con él. Menos Adictas al Estudio de la Cartografía, las Generaciones Siguientes entendieron que ese dilatado Mapa era *Inútil* y no sin Impiedad lo entregaron a las Inclemencias del Sol y los Inviernos. En los Desiertos del Oeste perduran despedazadas Ruinas del Mapa, habitadas por Animales y por Mendigos; en todo el País no hay otra reliquia de las Disciplinas Geográficas."

Jorge Luis Borges, Del Rigor en la Ciencia