

# ประเมินผลการตีความข้อมูล

## Evaluating Data Interpretation

หลีกเลี่ยงความผิดพลาดที่เกิดขึ้นบ่อยในการอ่านข้อมูล  
Avoiding common mistakes in reading your data

# ข้อผิดพลาดที่พบบ่อยในการตีความข้อมูล (Common Data Interpretation Mistakes)

---

- การเลือกข้อมูลตัวแทนประชากรกลุ่มตัวอย่าง (Proxy data) ที่น้อยไปและไม่ถูกต้อง
- การเปรียบเทียบสิ่งต่าง ๆ ที่ไม่เหมือนกัน
- ความสัมพันธ์  $\neq$  สาเหตุ
- ช่วงเวลาที่สั้นไม่สามารถนำไปวิเคราะห์หรือคาดการณ์แนวโน้มของข้อมูลในอนาคตได้
- ตรรกะวิบัติ (The Ecological Fallacy) - การให้อคติของคุณมีผลต่อการวิเคราะห์ข้อมูล

ข้อมูลตัวแทนประชากรกลุ่มตัวอย่าง (Proxy data) ที่ไม่เหมาะสม

## Choosing Bad Proxy Data

## Preventing the bite

The proportion of ownership of insecticide-treated has stagnated since 2007.



©Internews

School going children have higher prevalence of malaria than pre-schoolers.

In malaria-endemic areas, there is one insecticide-treated net to five Kenyans.

**1** insecticide-treated net is recommended for every five people.



2 out of 5 children under five years in the malaria-endemic Nyanza and Coast regions do not sleep under insecticide-treated nets.



3 JUNE 2015

### Liberia: Govt Warns Against Misuse of Mosquito Nets

Health • Liberia • Malaria • West Africa



The Ministry of Health and Social Welfare through the National Malaria Control Program is calling on Liberians to desist from misusing donated mosquito nets provided by the Government of Liberia.

Addressing a news conference Tuesday, National Malaria Control Program Manager Oliver J. Pratt said there are reports of individuals and communities selling and misusing the nets for football goal posts, bathing sponges and for fishing purposes, among others.

# ตัวอย่าง: ข้อมูลพรีอิกชีที่ไม่เหมาะสม (Examples of bad proxy data)

---

- รายงานคดีข่มขืน สำหรับความถี่ของคดีข่มขืน
- การจับกุมพ่อค้ายา สำหรับความถี่ในการใช้ยา
- รายงานความรุนแรงในครอบครัว สำหรับกรณีความรุนแรงในครอบครัว
- ความนิยมของผลิตภัณฑ์ เพื่อประสิทธิภาพของผลิตภัณฑ์นั้น

# ตัวอย่าง: ผลสำรวจค่าตอบแทน ลูกจ้างชาวต่างชาติในบ้าน

## FOREIGN DOMESTIC WORKER SALARY SURVEY

Middle East Edition, 2017

### AVERAGE SALARY OFFERED ACROSS THE REGION



AED1,812  
(US\$494)



QAR1,662  
(US\$450)



SAR1,638  
(US\$442)



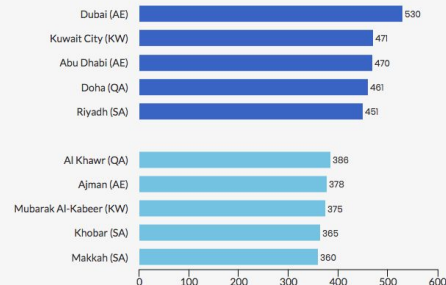
KWD126  
(US\$419)

### AVERAGE PAY BY CITY (US\$)



**US\$530**

EMPLOYERS IN DUBAI  
OFFER THE BEST SALARY



**US\$954**  
(AED3,500)

The highest salary offered by  
an employer in Abu Dhabi



**US\$813**  
(QAR3,000)

The highest salary offered by  
3 employers in Doha

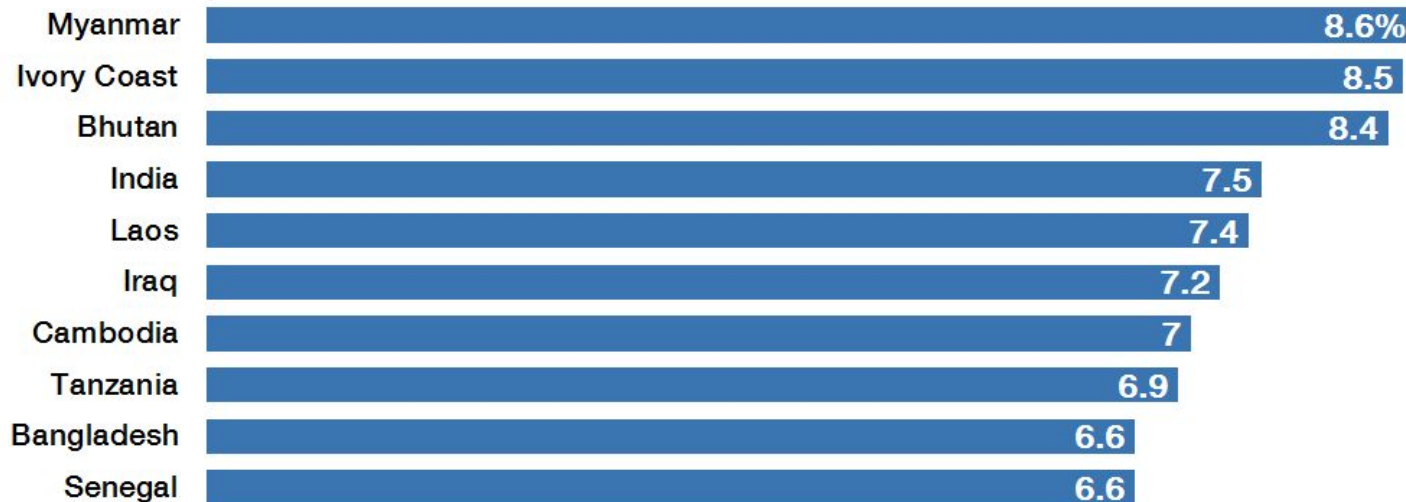
การเปรียบเทียบสิ่งต่าง ๆ ที่ไม่เหมือนกัน

**Comparing things that are not alike**

# ตัวอย่าง: การจัดอันดับประเทศที่มีเศรษฐกิจเติบโตเร็วที่สุดในโลก

## These are the world's fastest growing economies

Projected percentage real GDP growth, 2016



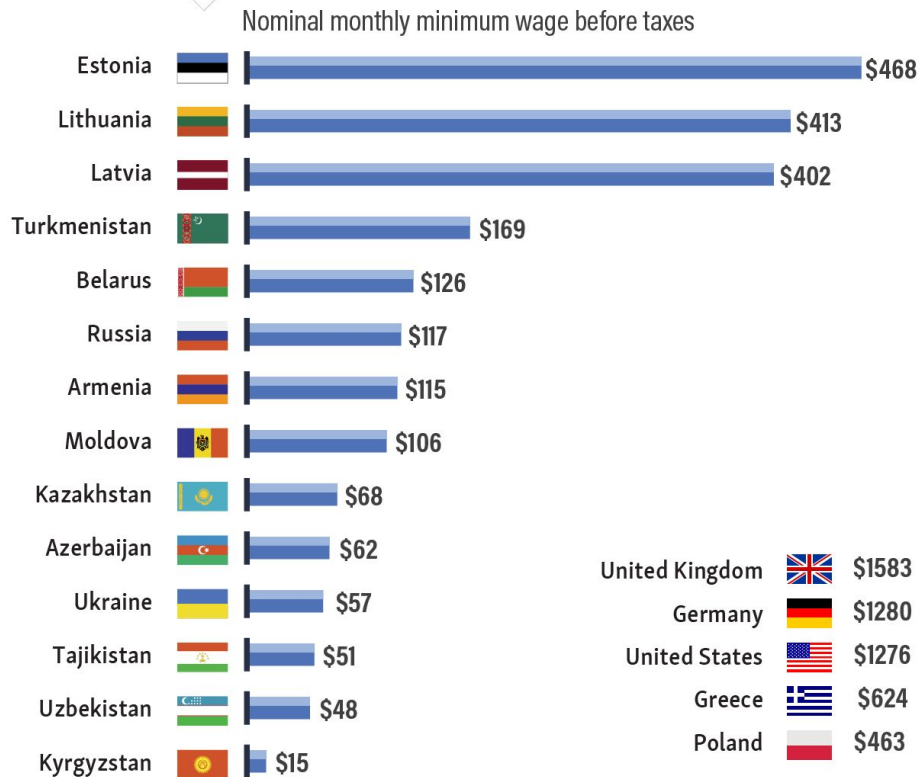
Source: IMF World Economic Outlook April 2016



ตัวอย่าง: เปรียบเทียบค่า  
แรงขั้นต่ำของกลุ่ม  
ประเทศภายหลังการล่ม  
สลายของสหภาพโซเวียต

## MINIMUM WAGE IN POST-SOVIET COUNTRIES

Which post-Soviet countries have the highest minimum wages?



Source: Official sources, EU Eurofound foundation

Notes: Figures based on official exchange rates on November 14, 2016

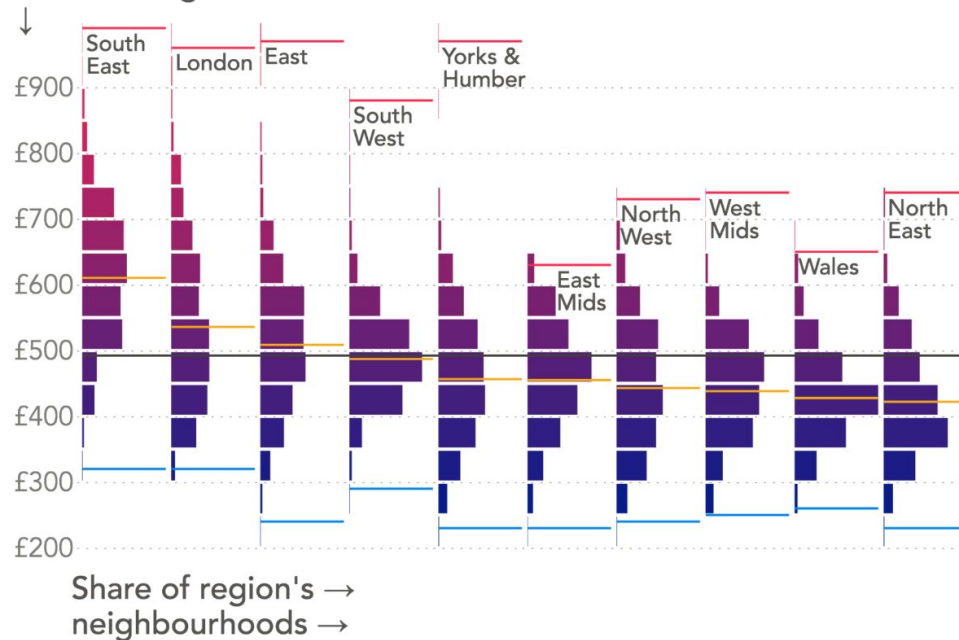
# เมื่อมีการเปรียบเทียบ... / When making comparisons...

---

- คุณมีเหตุผลที่ขับเคลื่อนด้วยข้อมูลสำหรับการเปรียบเทียบที่กำลังทำหรือไม่? (เช่น ประเทศที่ต้องการเปรียบเทียบมีสถานะทางสังคมและเศรษฐกิจที่คล้ายคลึงกันโดยวัดจากรายได้ต่อหัว หรืออายุขัยเป็นต้น)
- คุณควบคุมปัจจัยที่อาจจะทำให้การเปรียบเทียบไม่เป็นธรรมหรือไม่?
- คุณมองแนวโน้มของข้อมูลเมื่อเปรียบเทียบช่วงเวลาหรือไม่?

# ตัวอย่าง: รายได้ครัวเรือนต่อสัปดาห์ หลังหักค่าใช้จ่ายในบ้าน

Weekly household income after housing costs



ความสัมพันธ์ ≠ สาเหตุ

**Confusing Correlation and Causation**

# ความแตกต่างคือ? / What's the difference?

---

**ความสัมพันธ์ (Correlation)** เป็นความสัมพันธ์ร่วมกันของสองสิ่งหรือมากกว่านั้น หรือกล่าวได้ว่า ความสัมพันธ์ แสดงถึงค่าของตัวแปรสองตัวที่ขึ้นหรือลงในรูปแบบที่คล้ายกัน

ถือเป็นอุปายของความสัมพันธ์ เพราะยากที่จะรู้ว่าตัวแปรใดมีอิทธิพลต่ออีกหนึ่งตัวแปร หรืออีกคนหนึ่ง หากมีความสัมพันธ์ใด ๆ เกิดขึ้น

เมื่อเขียนพาดหัวเรื่อง มันเป็นไปได้ที่จะแสดงความสัมพันธ์โดยตรงระหว่างสองตัวแปร

ตัวอย่าง: ความสัมพันธ์  
ระหว่างการบริโภคช็อค  
โกแลตรายปี และ  
จำนวนรางวัลโน  
เบลต่อประชากร 10  
ล้านคน

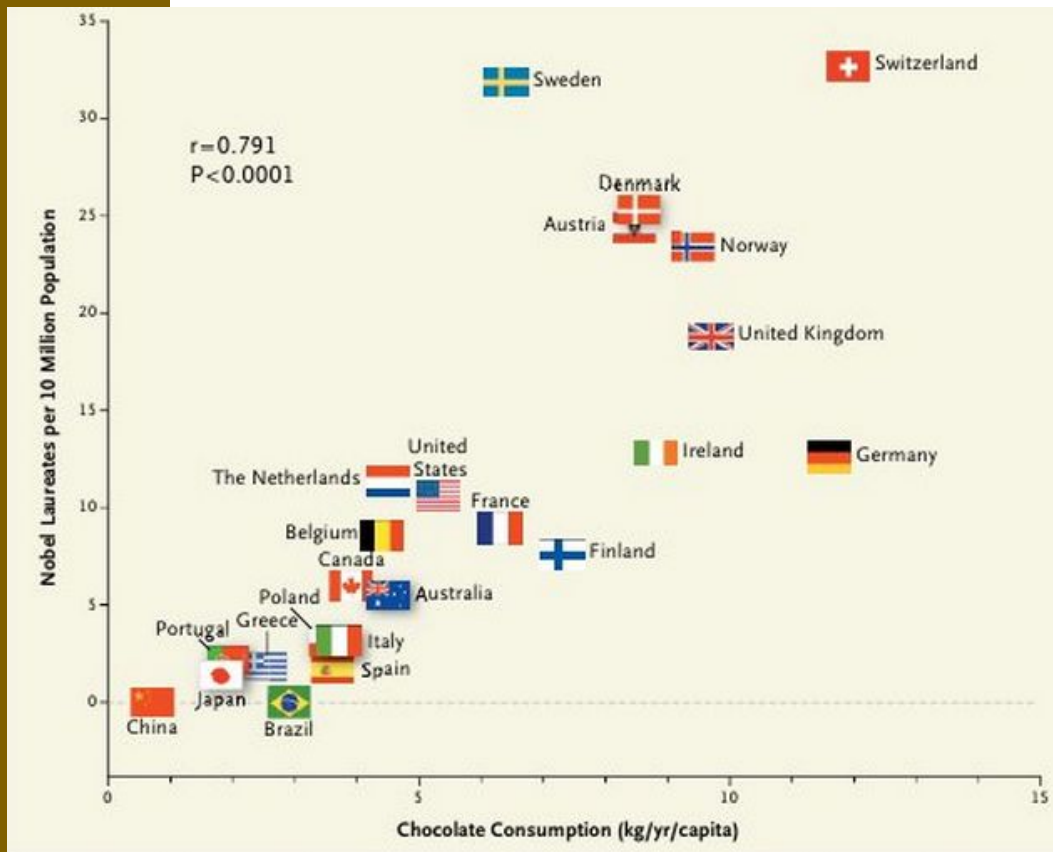


Figure 1. Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.

# รูปแบบความสัมพันธ์

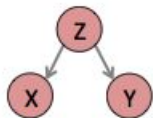
## How correlation happens



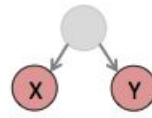
X causes Y



Y causes X



Z causes X and Y



hidden variable causes X and Y



random chance!



if you have a gun, you're going to use it



if it's a dangerous neighborhood, you'll buy a gun



the correlation is due to chance



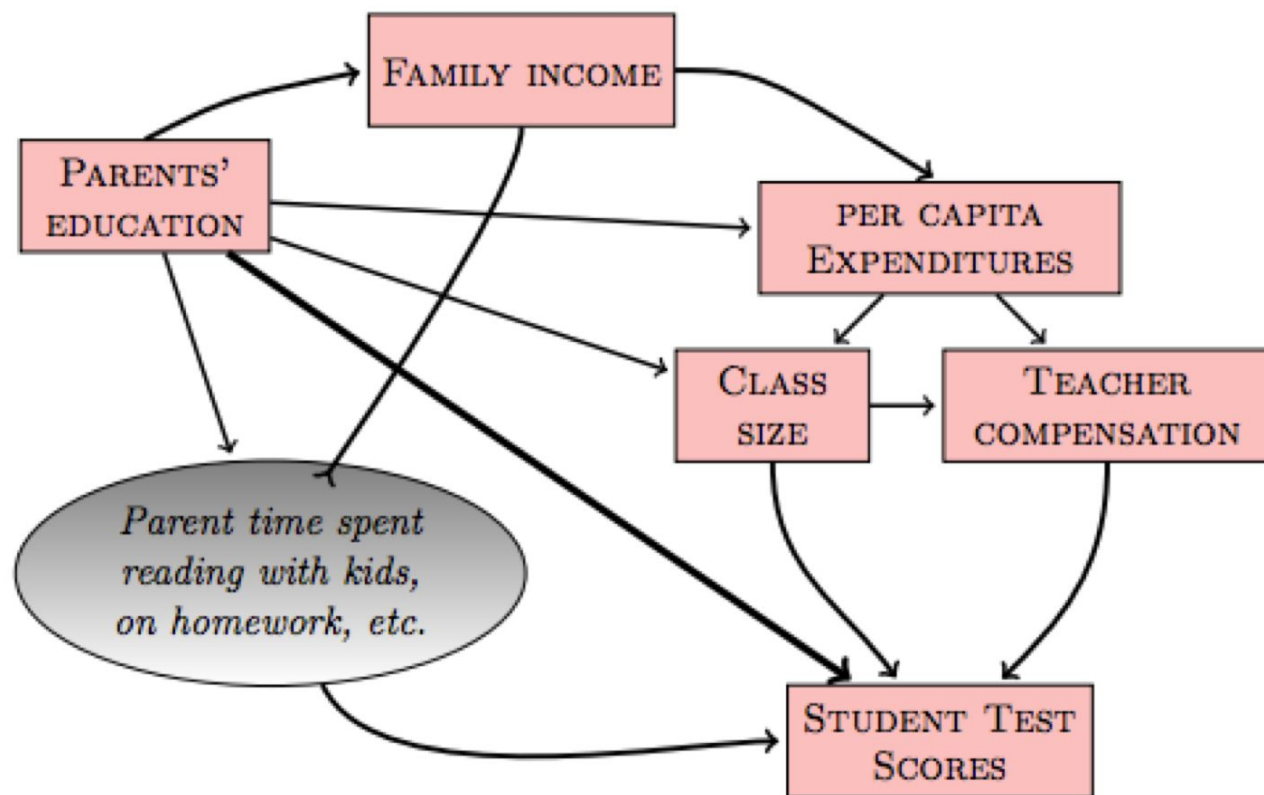


Figure 17.9: A hypothetical causal network relating student test scores to various educational policies and social conditions.

# How to keep a girl in school? Give her a toilet

By [Agencies](#) For Citizen Digital

🕒 Published on: September 8, 2015 06:59 (EAT)



ใช้ข้อมูลช่วงเวลาน้อยเกินไปสำหรับคาดการณ์แนวโน้มของข้อมูล

**Finding trends in too little data**

# Regression to the mean

---

แปลตรงตัวคือ การวิเคราะห์การถดถอย เพื่อหาค่าเฉลี่ย

Regression analysis เป็นการวิเคราะห์ถึงความสัมพันธ์ของตัวแปรอิสระที่เป็นเหตุ และตัวแปรตามซึ่งเป็นผลที่เกิดจากเหตุ หากตัวแปร 2 ตัวนั้น ในความเป็นจริงไม่มีความสัมพันธ์กันเชิงเหตุและผล เราจะใช้ Regression Analysis ไม่ได้

ถือเป็นเครื่องมือทางสถิติที่มีการประยุกต์ใช้ในการประมวลผลข้อมูลในงานวิจัยค่อนข้างมาก

# Regression to the mean

---

Q: ปริมาณผู้เสียชีวิตจากท้องถนนในเมือง A เท่ากับ 8 คนปีนี้ เทียบกับ 2 คนเมื่อปีก่อน และ 1 คนในปีก่อนหน้า ถามว่าเมืองของคุณอันตรายหรือไม่? คุณคิดว่าจะเกิดอะไรขึ้นในปีหน้า

Traffic deaths in your city: 8 persons this year, compared to only 2 last year and 1 the year before. Has your city become more dangerous? What do you think will happen next year?

Q. เปิดเครื่องปรับอากาศ วันละ 5 ชั่วโมง เดือนนี้จ่ายค่าไฟ 560 บาท ถ้าอยากจ่ายเดือนละ 450 บาท หรือน้อยกว่าควรเปิดวันละกี่ชั่วโมง (อุปกรณ์อย่างอื่นใช้เหมือนปกติ)

You turn on the a/c for 5 hours a day. This month you pay 560 Baht. If you want to pay 450 Baht or less, how many hours should you turn on the a/c (other home appliances remain the same)?

# Regression to the mean



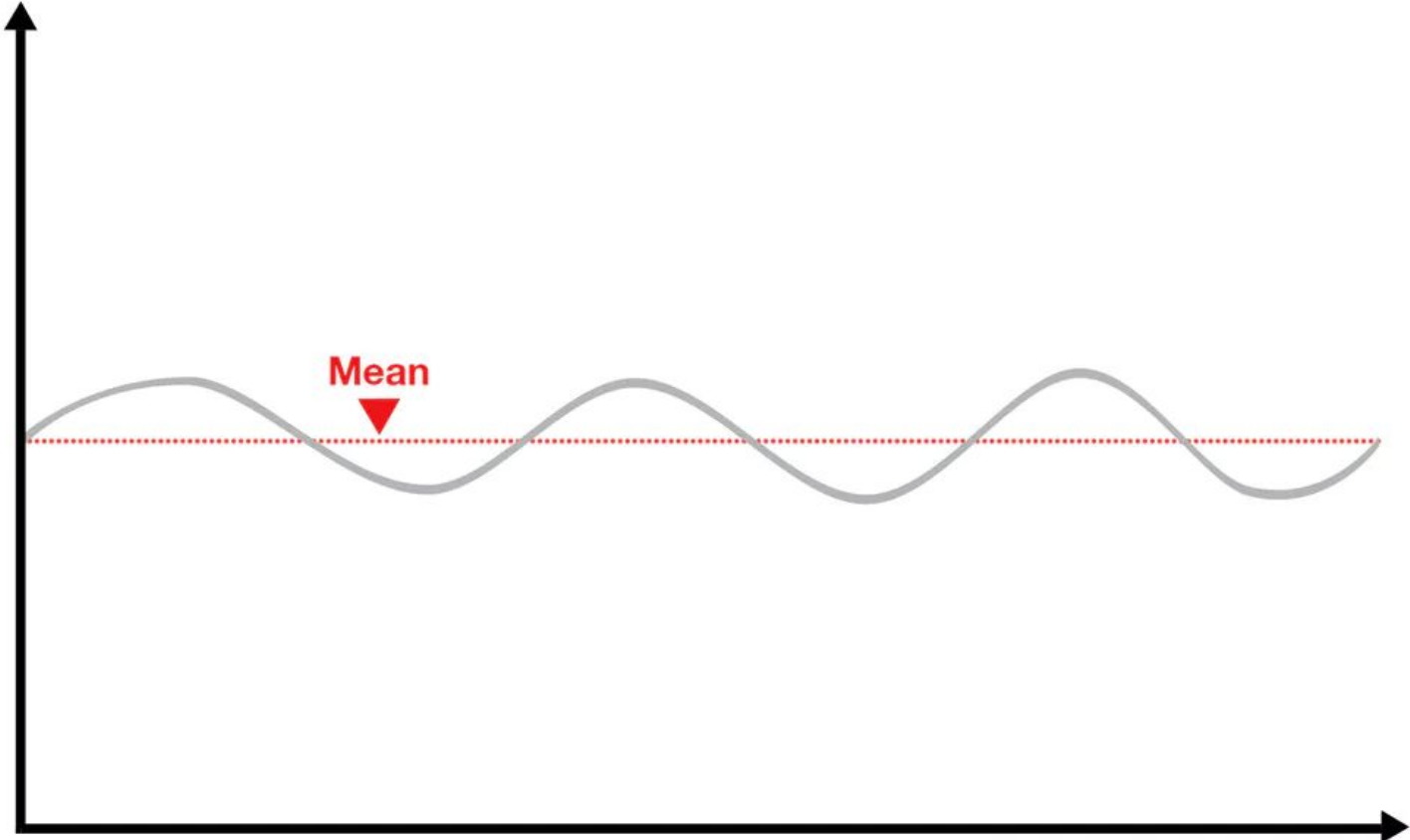
CRIME

## SLO saw “unprecedented” increase in rape, violent crimes in 2018, according to police

SYDNEY BRANDT - 2 HOURS AGO



# Regression/reversion to the mean





# Regression to the mean: the good news / ข้อดี

---

- เพียงเพราะคุณไม่มีข้อมูลสำหรับปีที่ผ่านมา ไม่ได้แปลว่าข้อมูลของคุณ "เก่า" เสมอ
- ย้อนดูข้อมูล 10 ปีที่ผ่านมา: ข้อมูลตัวเลขมีความสอดคล้องสม่ำเสมอ โดยยกเว้นข้อผิดพลาดหรือไม่?
- การใช้ข้อมูลแนวโน้มในอดีตย่อมจะดีกว่าการกำหนดให้ได้ข้อมูล "ใหม่ล่าสุด"

ตรรกะวิบัติ

**The Ecological Fallacy**

## ตรรกะวิบัติ (The Ecological Fallacy)

---

# No Place for Foreigners

Why Hanna is invited to view the apartment and Ismail is not

*If you have a foreign name and you are looking for an apartment, you will significantly less likely to be contacted by landlords than a German applicant. In particular, people with Arab and Turkish names are severely discriminated against. This is shown by a large-scale experiment conducted by the data journalists of Bayerischer Rundfunk and SPIEGEL.*

# The Ecological Fallacy

---



**Hanna Berg**

27 years old, single, works in the marketing sector. Hanna speaks and writes flawless German.



**Ismail Hamed**

27 years old, single, works in the marketing sector. Ismail speaks and writes flawless German.

# The Ecological Fallacy

---

ข้อมูลของแต่ละบุคคลสามารถรวมกันเพื่อบอกเล่าเรื่องราวเกี่ยวกับกลุ่มได้

ข้อมูลของกลุ่มไม่สามารถจำแนกเพื่อบอกเล่าเรื่องราวรายบุคคลในกลุ่มนั้นได้

**อย่าอนุมานเรื่องของคนจากข้อมูลสถิติกลุ่ม  
(Don't infer things about individuals when you only have group statistics)**

# ตัวอย่าง (Sample): The Ecological Fallacy

---

- จากอัตราการทำแท้งที่ลดลง หนังสือพิมพ์มีสมมติฐานเกี่ยวกับทางเลือกอื่น ๆ ของผู้หญิง
- ข้อสรุปที่ผิดพลาดว่านักเรียนทุกคนในห้องที่มีคะแนนสูงสุดระดับประเทศในวิชาคณิตศาสตร์จะต้องมีความพิเศษทางคณิตศาสตร์
- รายงานกล่าวยืนยันอย่างไม่ถูกต้องว่า การเพิ่มขึ้นของใบสั่งยาากล่อมประสาททั่วประเทศ บ่งชี้ถึงวิกฤตของภาวะซึมเศร้าในหมู่ประชาชน
- การเหมารวมว่าผู้หญิงทุกคนเป็นคนขับรถที่แย่

-Heather Krause

## FB: Anti Fallacy ต่อด้านตรรกะวิบัติ

### Ecological Fallacy

#ด่วนสรุปจากหลักฐานแวดล้อม

เป็นการสรุปหรืออนุมานว่า  
คนหรือสิ่งหนึ่งจะเป็นหรือประพฤติ  
เหมือนกับกลุ่มของคนหรือสิ่งนั้นๆ

เช่น การสรุปหรือตัดสิน  
บุคคลหรือสิ่งของจาก ค่าเฉลี่ย

#ต่อด้านตรรกะวิบัติ

ตรรกะวิบัติ #14

นักเรียนคนนี้มีเกรดเฉลี่ย 1.72

∴ นักเรียนคนนี้ไม่เก่งทฤษฎี

“เธอมาจากมหาวิทยาลัยนี้  
แสดงว่าเธออ่อนกว่าคนอื่นสินะ”

ทะเลทรายนี้มีอุณหภูมิเฉลี่ย 24°C

24°C เป็นอุณหภูมิที่อบอุ่นสบาย

∴ ทะเลทรายมีสภาพอากาศอบอุ่น

“คุณมาจากจังหวัดนี้  
แสดงว่าคุณจน”

#ต่อด้านตรรกะวิบัติ

ตรรกะวิบัติ #14

# ข้อผิดพลาดที่พบบ่อยในการตีความข้อมูล (Common Data Interpretation Mistakes)

---

- การเลือกข้อมูลตัวแทนประชากรกลุ่มตัวอย่าง (Proxy data) ที่น้อยไปและไม่ถูกต้อง
- การเปรียบเทียบสิ่งต่าง ๆ ที่ไม่เหมือนกัน
- ความสัมพันธ์  $\neq$  สาเหตุ
- ช่วงเวลาที่สั้นไม่สามารถนำไปวิเคราะห์หรือคาดการณ์แนวโน้มของข้อมูลในอนาคตได้
- ตรรกะวิบัติ (The Ecological Fallacy) - การให้อคติของคุณมีผลต่อการวิเคราะห์ข้อมูล



# Thank You



This lesson was adapted from the World Bank's Introduction to Data Literacy training manual by Eva Constantaras, and adapted by Yan Naung Oak, Open Development Cambodia and Open Development Initiative, and is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). For full terms of use, see [here](#).