

การใช้ข้อมูลขนาดใหญ่เพื่อ การวิจัยและการสอน

ศ.ดร.ติเรก ปัทมสิริวัฒน์

วันที่ ๑๖ มกราคม ๒๕๖๑

มหาวิทยาลัยนเรศวร

การวิจัยเพื่อสร้างองค์ความรู้และการสอนระดับแนวหน้า

1. สอนดีต้องมีวิจัย
2. การวิจัยเพื่อตอบโจทย์ประเทศ ยุทธศาสตร์ประเทศไทย 4.0 โดยไม่ทอดทิ้งกัน
3. มหาวิทยาลัยมีบทบาทการวิจัย การเสนอแนะนโยบาย ให้บริการกับจังหวัด และหน่วยงานที่ขอรับบริการวิชาการ
4. ในที่นี้ขออภิปราย การนำข้อมูลขนาดใหญ่มาสอน/วิจัยเพื่อก้าวสู่แนวหน้าและเพิ่มบทบาทด้านนโยบายสาธารณะ

ข้อมูลขนาดใหญ่ในยุคปัจจุบัน

1. การบันทึกข้อมูลขนาดใหญ่ เป็นแนวโน้มแห่งยุค ขอบคลุมความก้าวหน้าทางเทคโนโลยี คอมพิวเตอร์ อินเทอร์เน็ตและหน่วยงานที่บันทึกข้อมูลขนาดใหญ่
2. ตัวอย่างของข้อมูลขนาดใหญ่
 - การซื้อขายที่ดิน จดทะเบียนกับกรมที่ดิน
 - แรงงาน/สถานประกอบการในระบบประกันสังคม เงินสมทบเข้ากองทุนประกันสังคม หรือ กบข
 - บัญชีเงินฝากในระบบธนาคาร
 - การประกันชีวิต เบี้ยประกัน วงเงินทำประกัน

ข้อมูลขนาดใหญ่ในยุคปัจจุบัน

1. ตัวอย่างของข้อมูลขนาดใหญ่

- การรักษาพยาบาล ภายใต้ 3 ระบบ ข้าราชการ ประกันสังคม ประกันสุขภาพถ้วนหน้า
- การใช้บริการฉุกเฉิน **ems=emergency medical service**
- การจัดสรรงบประมาณรายจังหวัด
- กองทุนนอกงบประมาณ
- สหกรณ์ออมทรัพย์ ทรัพย์สินหนี้สิน จำนวนสมาชิก เงินฝาก

การเข้าถึงข้อมูลและนำมาวิจัย

ข้อมูลขนาดใหญ่ จัดทำโดยหน่วยงานของรัฐ บันทึกเป็นรายครึ่ง – รายบุคคล – รายแปลง – ซึ่งอาจจะเป็นความลับ การเข้าถึงข้อมูลขึ้นอยู่กับหน่วยงาน

แต่ข้อมูลที่ประมวลเป็น รายจังหวัด หรือ รายโซน หน่วยงานมักเปิดเผย ซึ่งมีคุณสมบัติประโยชน์อย่างมากต่อนักวิจัย

เช่น ข้อมูลซื้อขายที่ดิน สินเชื่อ เงินฝากธนาคาร การทำประกันชีวิต .. ประมวลเป็นจังหวัด ข้อมูลรายปี/รายเดือน ...สามารถเรียบเรียงเป็น **panel data** เพื่อการวิเคราะห์

นำข้อมูลขนาดใหญ่เพื่อวิเคราะห์พลวัตของเศรษฐกิจจังหวัด

1. แต่ข้อมูลขนาดใหญ่ที่กล่าวมานี้ สามารถประมวลเป็นระดับ จังหวัด .. มหาวิทยาลัย เข้าถึงได้ และนำมาสร้างสรรค์งานวิจัย/และใช้เพื่อการเรียนการสอน ได้เป็นอย่างดี ทั้งนี้ ต้องประมวลเป็น **panel data (77 จังหวัด x 20 ปี)**
2. คุณประโยชน์ข้อมูลแบบนี้ หนึ่ง วิเคราะห์พลวัตของจังหวัดได้ สอง วิเคราะห์ความแตกต่าง/เหลื่อมล้ำได้ สาม สนับสนุนการทำงานของหน่วยงานระดับจังหวัด และ ยุทธศาสตร์เชิงพื้นที่ **area-based approach** สี่ การเรียนการสอนที่น่าตื่นเต้น เห็น ปัญหาในระดับพื้นที่ **problem-based learning** ห้า สะท้อนความก้าวหน้าทาง เทคนิคและวิชาการสมัยใหม่

แบบจำลองพลวัตของจังหวัด

1. **Provincial growth model** และปัจจัยที่เกื้อหนุน
2. ทุนและสินเชื่อ สามารถใช้ **bank credit** เป็นตัวแทน
3. แรงงานในระบบประกันสังคม
4. สถานประกอบการในระบบประกันสังคม
5. การซื้อขายที่ดิน เป็น **proxy** ของการคาดการณ์อนาคต **speculative**
6. รายจ่ายภาครัฐ — รายจ่ายขององค์กรปกครองส่วนท้องถิ่น สะท้อนตัวแปรนโยบายการคลัง

สถิติและเศรษฐมิติเพื่อวิเคราะห์พลวัต และ ความแตกต่าง

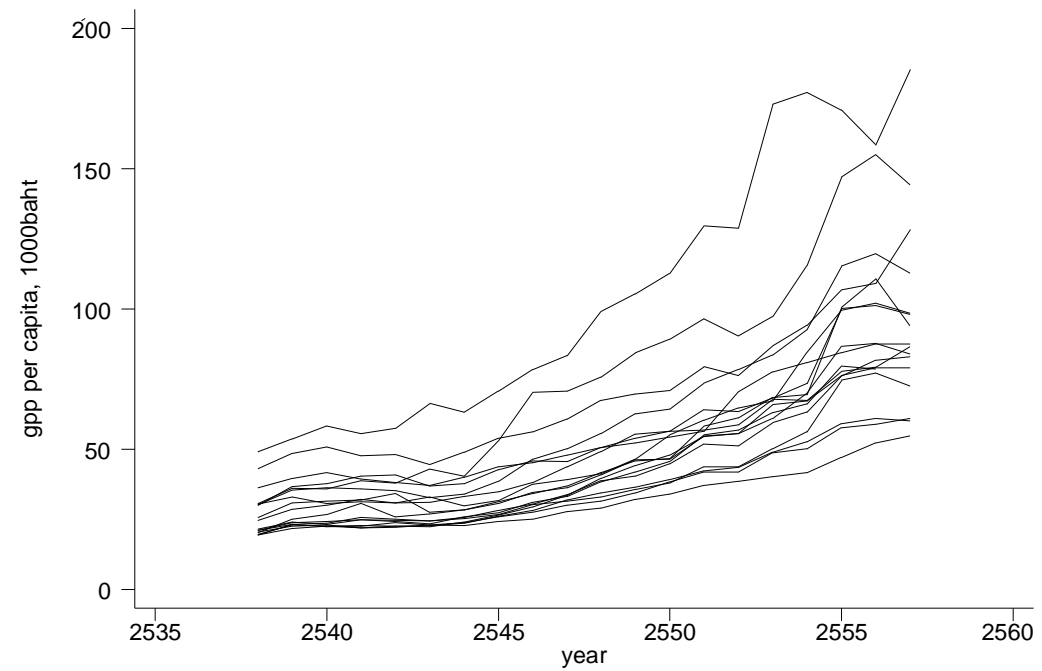
1. ความก้าวหน้าด้านเทคนิคการคำนวณ **software** ช่วยให้แบบจำลองที่ซับซ้อน กลายเป็นเรื่องง่าย การใช้กราฟแสดงการเปลี่ยนแปลง ดัชนีวัดความเหลื่อมล้ำมิติพื้นที่
2. ตัวอย่าง **provincial growth panel : northern region (17 จังหวัด 21 ปี ข้อมูล 1995-2015)**
3. การวิเคราะห์เชิงระบบ **system analysis**, ปัจจัยนำเข้า ผลกระทบ วัดความแตกต่าง/เหลื่อมล้ำ ... ช่วยให้นักวิจัยวิพากษ์นโยบายของรัฐ หรือเสนอแนะการเปลี่ยนแปลงของจังหวัด **area-based approach**

ผลดี..อานิสงส์ของการใช้ข้อมูลขนาดใหญ่ (ผ่านการปรุง)

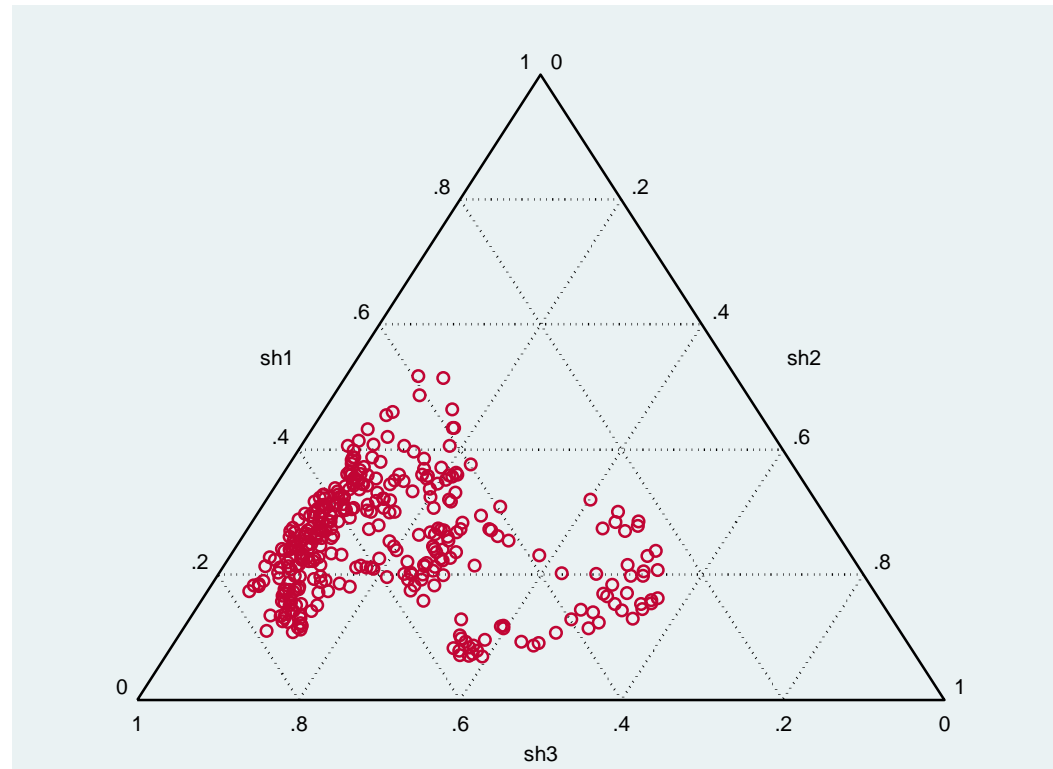
1. ความก้าวหน้าวิทยาการ การประยุกต์ (กับสภาพปัญหาหรือมีกรณีศึกษา) การเสนอแนะนโยบาย การวิพากษ์นโยบายระดับจังหวัด/พื้นที่
2. การเรียนการสอน สนุก น่าตื่นเต้น .. ก้าวเข้าหา **frontier**
3. การเชื่อมโยงมหาวิทยาลัย กับ หน่วยงานภาคปฏิบัติ จังหวัด อปท. ชุมชน สหกรณ์

ตัวอย่างแบบจำลอง

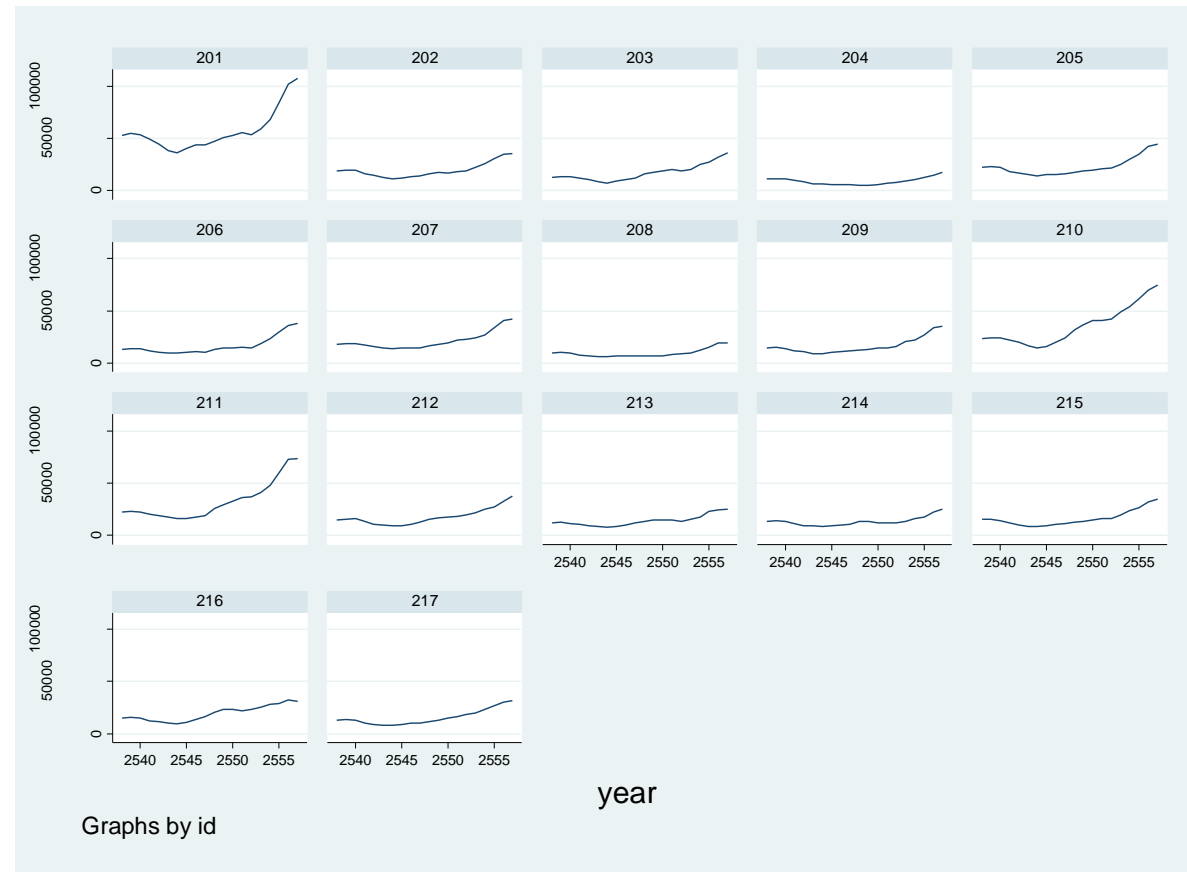
1. Provincial growth



PRODUCTIVE STRUCTURE: 17 NORTHERN PROVINCES



BANK LOAN PER CAPITA: 17 NORTHERN PROVINCES



สถิติทางเศรษฐกิจของ จ.เชียงใหม่ (ID=201)

year	estb	insure1	loan	pop1000	gpp
2550	11966	176512	88603	1679.141	119146.6
2551	12171	174380	94562	1698.522	135019.2
2552	12555	175878	92297	1717.315	130955.8
2553	13232	178108	102439	1735.48	150801.9
2554	13530	181747	118450	1732.48	163230
2555	13634	185208	144980	1729.38	184886.4
2556	13972	190381	176267	1725.58	188323.3
2557	14458	197100	185209	1721.48	221215.6
Total	13189.75	182414.3	125350.9	1717.422	161697.3

แบบจำลองท้องถิ่นไทย

องค์กรปกครองส่วนท้องถิ่น **7,853** แห่ง เฉพาะเทศบาล-อบต. **7775** แห่ง
ประเภทข้อมูล

- พื้นฐาน ประชากร พื้นที่ ความหนาแน่น
- รายได้ จำแนก รายได้จัดเก็บเอง ภาษีที่รัฐจัดเก็บให้ เงินอุดหนุน
- พนักงาน ข้าราชการ ลูกจ้าง
- รายจ่าย รายจ่ายบุคลากร รายจ่ายงบกลาง รายจ่ายดำเนินการ รายจ่ายการลงทุน รายจ่ายอุดหนุน

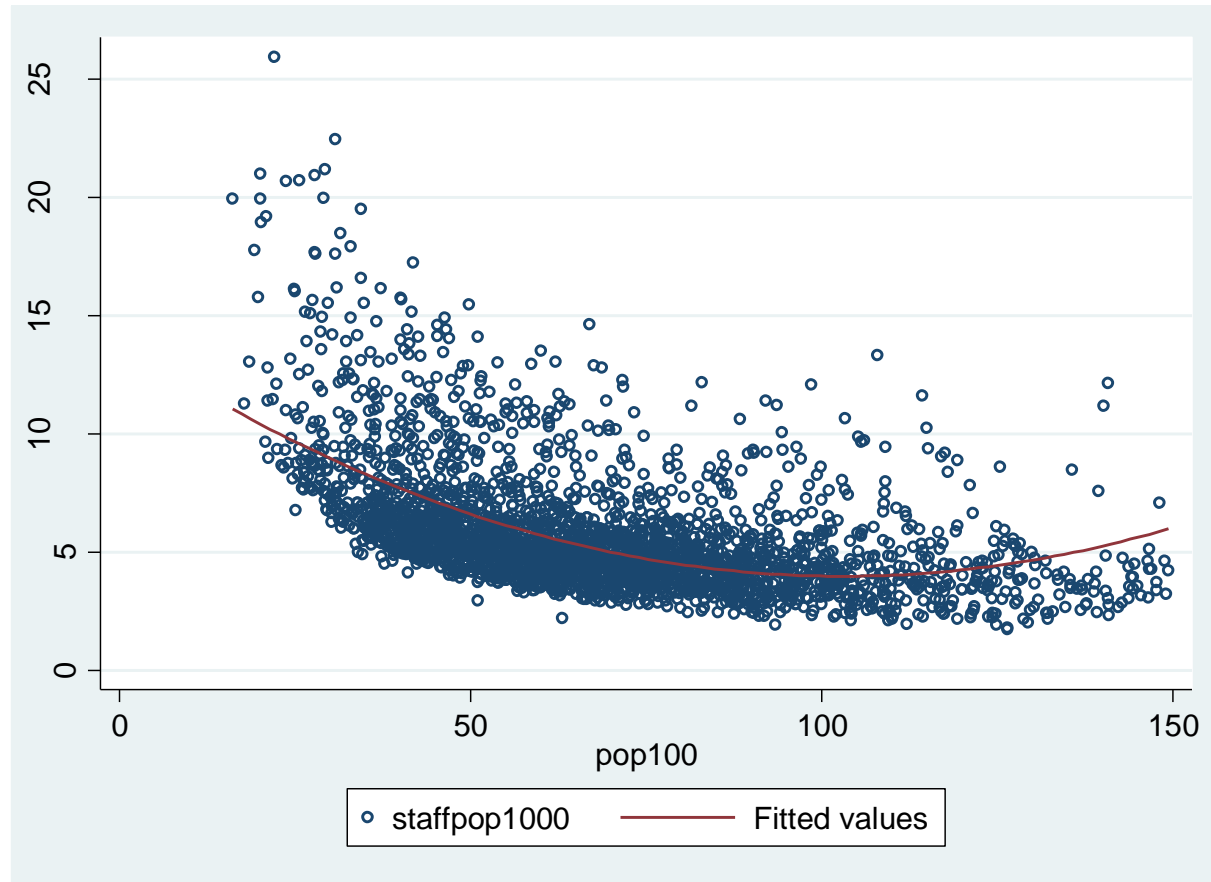
การวิเคราะห์การประหยัดจากขนาด ศักยภาพการควบรวมท้องถิ่น

ปัญหาคือ อปท. จำนวนหนึ่งขนาดเล็กเกินไป ประชากรน้อย พื้นที่เล็ก
ลักษณะ ตำบลเดียวกัน มีทั้งเทศบาลตำบล และ อบต. ลักษณะ ไข่มด
ขนาดเล็กเกินไป ไม่ประหยัดจากขนาด **staff/pop1000** สูง ใช้จ่าย
บุคลากรสูง

ข้อเสนอ (รัฐบาล สภาขับเคลื่อนการปฏิรูป) เสนอให้ ยก อบต. เป็นเทศบาล
ควบรวม อปท. ขนาดเล็ก เช่น ตำบลเดียวกันมีทั้งเทศบาลตำบล และ อบต.
ควรรควบรวม

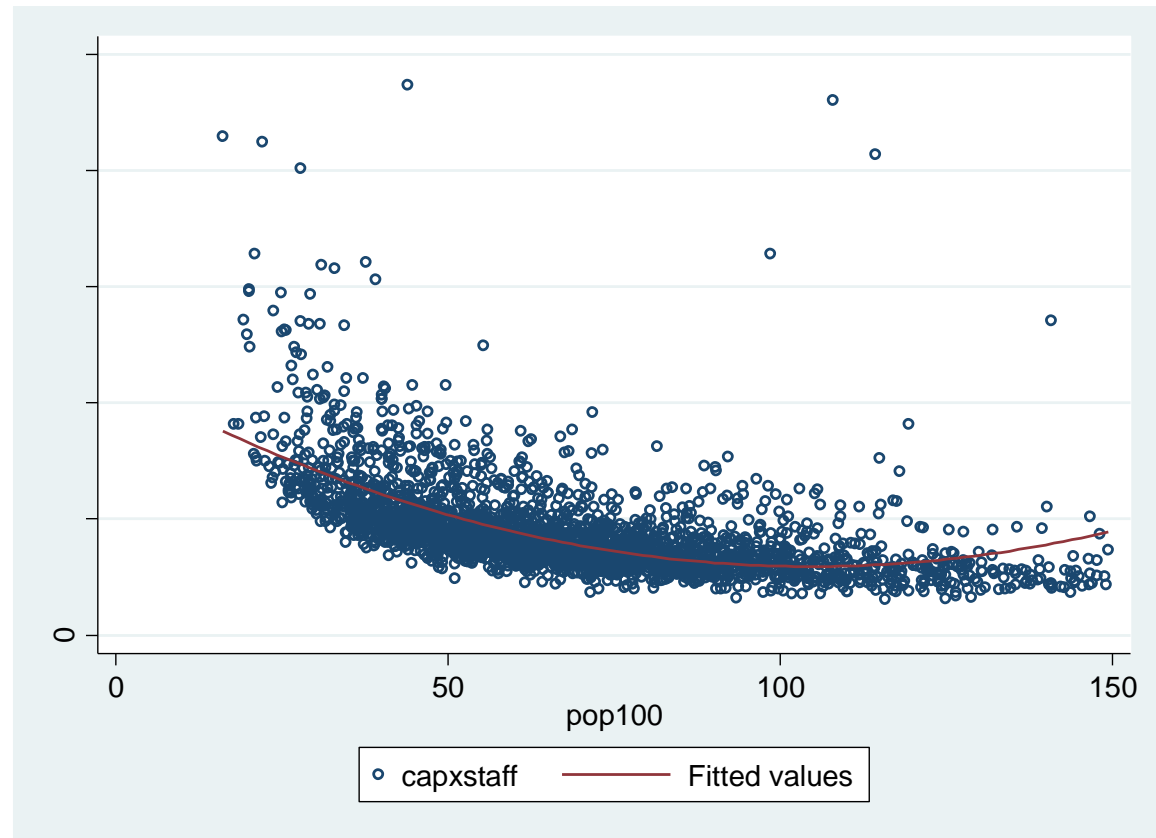
แต่การควบรวมมีอุปสรรค การเมืองท้องถิ่นไม่อยากจะเปลี่ยนแปลง/ปรับตัว

การประหยัดจากขนาดในเทศบาล อบต. หน่วยวัด STAFF / POP1000 กรณีศึกษาภาคเหนือ



การประหยัดจากขนาดในเทศบาล อบต.

หน่วยวัด รายจ่ายบุคลากรต่อประชากร กรณีศึกษาภาคเหนือ



ต่อจาก **BIG DATA** เสริมด้วยข้อมูลเชิงคุณภาพ **SMALL DATA**

กรณีศึกษา การรวบรวม อบต.วังเหนือ และ เทศบาลตำบลวังเหนือ จังหวัดลำปาง
ประกาศในราชกิจจานุเบกษา พุทธศักราช **2560** มีผลบังคับ **1 ธันวาคม 2560**

การเยี่ยมชมพื้นที่ เก็บข้อมูลเชิงคุณภาพ เข้าใจความเป็นมา รู้ปัญหา/อุปสรรค

ประเมินความเป็นไปได้ของการรวบรวม

นโยบายรวบรวม จะก้าวหน้าดี ควรจะมี ความรู้เสริม จากสถาบันวิชาการในท้องถิ่น การ
รวมหน่วยงานเข้าด้วยกัน มีค่าโสหุ้ย การปรับปรุงฐานข้อมูล ... เป็นบทบาทที่มหาวิทยาลัยใน
ภูมิภาคมีโอกาสแสดงบทบาทมาก มหาวิทยาลัยราชภัฏ

แต่ความเป็นจริง มีช่องว่างระหว่างคนมหาวิทยาลัย กับ หน่วยงานปฏิบัติจังหวัด/อำเภอ/
ตำบล

การข้ามพรมแดนศาสตร์

วรรณทอง จาก ศ. ดร.เจตนา นาควัชระ ความรู้ส่งทางซึ่งกันและกัน

การทำงาน (วิจัย) ข้ามศาสตร์ เกิดขึ้นได้เสมอ และเป็นของดี

วิเคราะห์ ระบบการแพทย์ฉุกเฉิน (**EMS**) ต้องการให้นักเศรษฐศาสตร์วิเคราะห์ **unit cost** เพื่อใช้ประกอบการจัดสรรงบประมาณ เราได้รับข้อมูล **EMS** เป็นรายครั้ง (หลายแสน/ล้านครั้ง) ทั่วประเทศ **demand side, supply side**

ตัวอย่างข้อมูล **EMS** ผู้ป่วยฉุกเฉิน อุบัติเหตุ หน่วยกู้ภัย โรงพยาบาล วิธีการรักษา ผลลัพธ์ เวลาที่ใช้เดินทาง

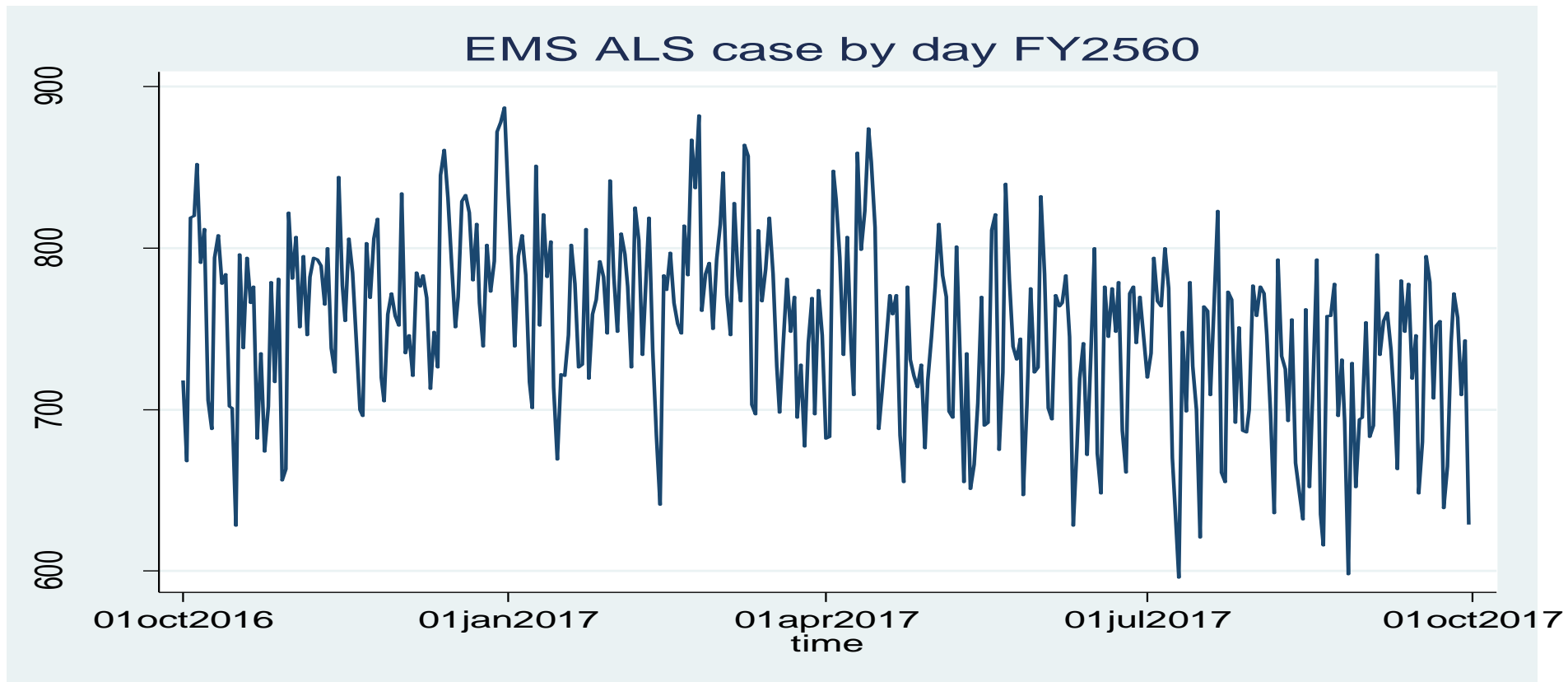
ระบบ **EMS** การลงทุนของรัฐ อบจ. เทศบาล มูลนิธิ ค่าตอบแทนแพทย์พยาบาล คนขับรถ

การข้ามพรมแดนศาสตร์

ตัวอย่างข้อมูล **EMS** รายบุคคล (ครั้ง) ระบุจังหวัด วันเวลา หน่วยปฏิบัติการ โรงพยาบาล
เวลา ระยะทาง วิธีการรักษา

Variable	Obs	Mean	Std. Dev.	Min	Max
pvcode	273,972	45.88283	23.6547	10	96
oper	273,120	1.470486	1.529286	1	8
case	273,972	1	0	1	1
case1	269,000	1	0	1	1
admit1	273,972	.4584958	.4982753	0	1
traffic	273,972	.2127918	.4092824	0	1
recover	273,972	.2100397	.4073373	0	1
dead	273,972	.0258019	.1585442	0	1

ข้อมูลการใช้บริการ EMS รายจังหวัดรายวัน



ความเหลื่อมล้ำด้านการคลัง/งบประมาณ

ฐานข้อมูลท้องถิ่น สามารถตอบวิเคราะห์ความเหลื่อมล้ำ ระหว่างจังหวัด หรือ ระหว่างเทศบาล / อบต.

name	caprev	self fin	capx staff	capxserv	name	caprev	self fin	capx staff	capxserv
chiang mai	7,630	6.2	1,945	2,635	phayao	7,144	3.6	1,989	2,420
chiang rai	7,308	4.1	1,817	2,383	phetchabun	6,554	3.3	1,576	2,311
kamphaeng phet	7,027	4.4	1,746	2,647	phichit	8,157	2.8	2,122	2,855
lampang	7,157	3.5	1,880	2,309	phitsanulok	7,665	5.4	1,964	2,447
lamphun	8,211	5.7	1,990	3,055	phrae	8,216	3.3	2,126	2,301
mae hong son	7,994	2.9	2,098	3,297	sukhothai	7,854	3.4	2,176	2,545
nakhon sawan	7,116	3.9	1,647	2,308	tak	7,113	4.2	1,805	2,855
nan	7,878	3.0	2,060	2,468	uthai thani	8,330	3.5	2,261	3,538
					uttaradit	9,063	4.0	2,392	2,969
					Total	7,589	4.1	1,941	2,596

การพัฒนาเป็น ฐานข้อมูล

จากข้อมูลดิบ เรานำมารวบรวมเป็น ฐานข้อมูลเพื่อการวิเคราะห์

การออกแบบข้อมูล **cross-sectional, time-series, panel data**

เช่น ข้อมูล **EMS** เป็นรายครั้ง รายวัน รายชั่วโมง ระบุพื้นที่ การปฏิบัติการ การรักษาพยาบาล

สามารถประมวลเป็นรายจังหวัด เพื่อการเปรียบเทียบ **cross-sectional**

เป็นรายวัน-เดือน-ปี **time-series**

หรือเป็น **panel data, $Y_{it}, i=\text{province}, \text{time}=\text{daily}$**

เครื่องมือ/โปรแกรม

ขอพูดถึง **STATA**

ซึ่งมีคุณลักษณะที่ดี มีเทคนิคการคำนวณที่ทันสมัย เป็นโปรแกรมที่ปรับปรุงตลอดเวลา มีแฟนคลับทั่วโลก และวารสาร **STATA Journal**

Estimate panel data

โปรแกรม **econometric evaluation model** การควบคุมตัวแปรอื่น ๆ ให้เท่ากัน เพื่อเปรียบเทียบ **treatment effect** เช่น **adjusted risk**

ตัวอย่างการวิจัยและสาธิตการนำเข้า-การวิเคราะห์

การซื้อขายที่ดิน

- วัดปริมาณซื้อขายที่ดิน เปรียบเทียบกับ จำนวนโฉนดที่ดิน รายจังหวัด
- ศึกษาการคลัง ภาษีการโอนที่ดิน ส่วนหนึ่งเป็นรายได้ให้เทศบาล และ อบต

การคลังของ กทม.

- **50** สำนักงานเขต จำนวน **20** ปีเศษ
- วัดความเหลื่อมล้ำการคลัง วัดความสามารถจัดเก็บรายได้

พลวัตการเติบโตของจังหวัด **76x21** ปี

- ศึกษาความเหลื่อมล้ำระหว่างจังหวัด และทดสอบข้อสันนิษฐาน การไล่กวดทางเศรษฐกิจ **catching up hypothesis**