

Data Visualization and Public Data Availability

Daniel Ray Lewis
Chulalongkorn University

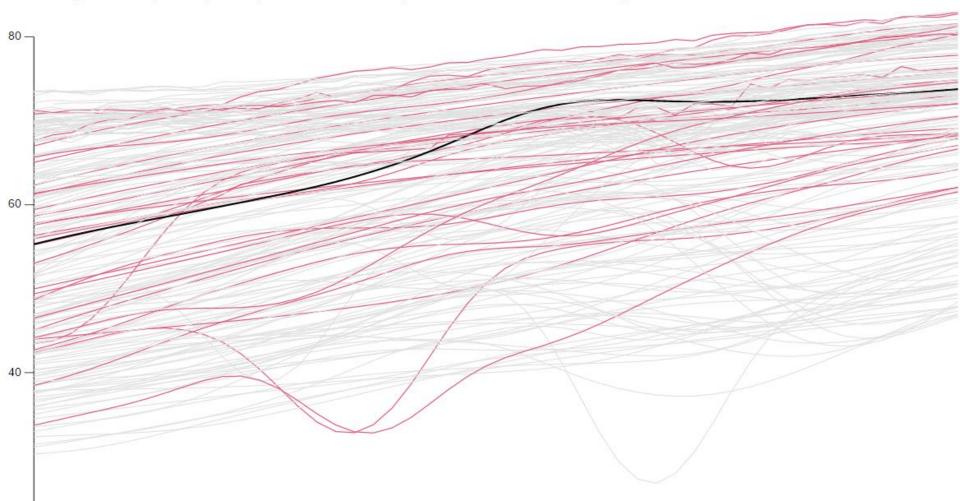
Life Expectancy

People are living longer around the world, some more so than others. Select a region (as defined by World Bank) below to compare, or roll over to the graph to highlight countries. Read more...

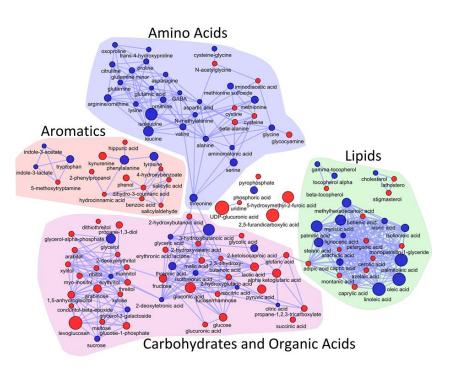
East Asia and Pacific South Asia Europe and Central Asia Middle East and North Africa Sub-Saharan Africa Latin America and Caribbean North America

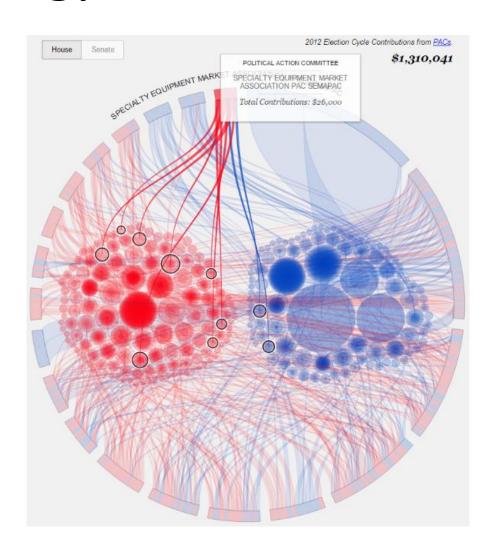
THAILAND

On average: a life expectancy of 55 years in 1960 and 74 years in 2009, an increase of 33 percent.

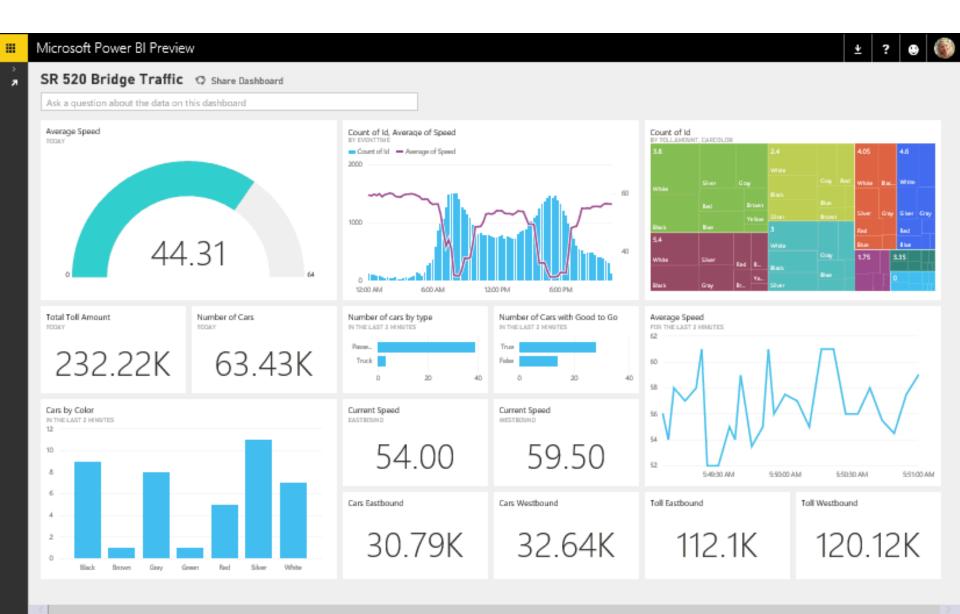


Clustering – Biology and Politics





Real time Dashboards



My Big Data Research

- THREE OBJECTIVES
- Objective 1: To determine Who receives the benefit from government policies (within a province or area, income range, profession, etc.)
- **Objective 2:** To determine the effect of government policies on regional GDP ECONOMIC Cost Benefit
- Objective 3: To determine the effects of government policies on Voting patterns – POLITICAL Cost-Benefit

Data Used

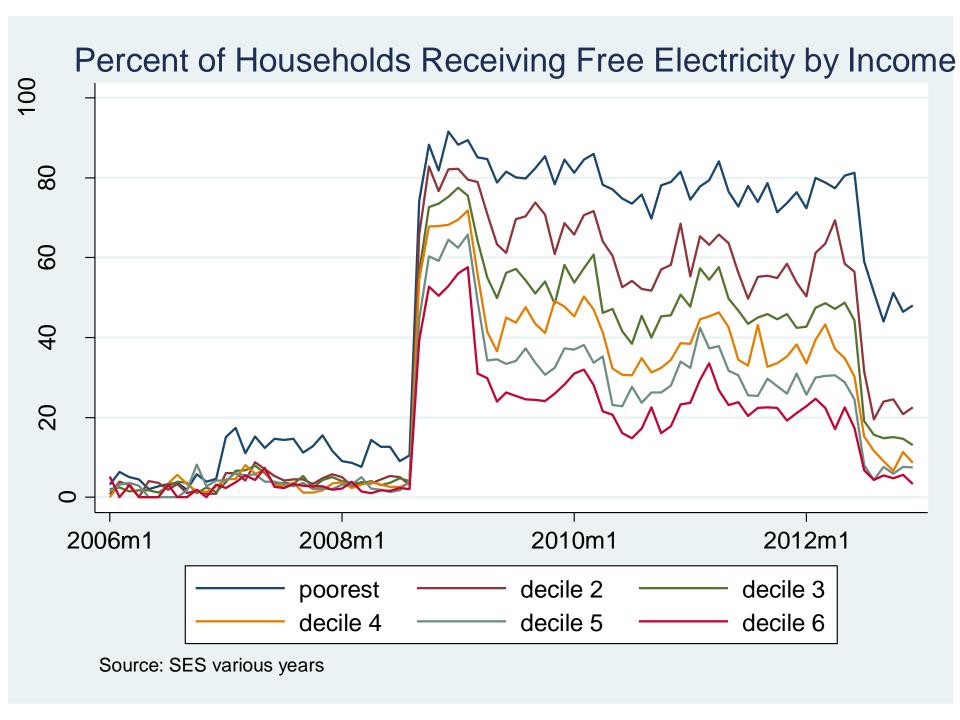
SES Data

- 2006 570 variables, 44,918 observations
- 2007 566 variables, 43,055 observations
- 2008–385 variables, 44,969 observations
- 2009 595 variables, 43,844 observations
- 2010 392 variables, 44,273 observations
- 2011 545 variables, 42,192 observations
- 2012 411 variables, 43,762 observations
- 2013 594 variables, 42,738 observations

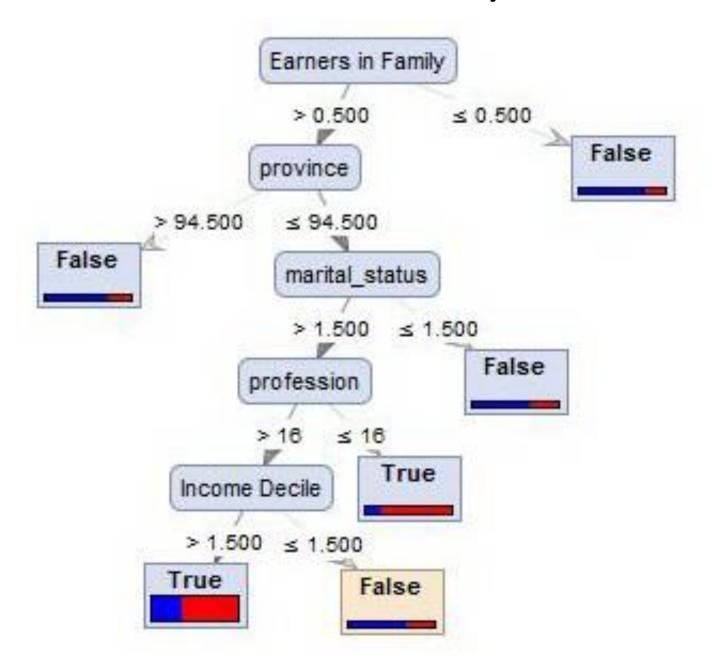
NESDB

- 19 years 77 provinces 23 subcategories
- Electoral Commission of Thailand
 - 2007 76 provinces 41 parties
 - 2011 77 provinces 40 parties

Altogether a bit over 20 million data points



Decision Tree – Who buys Condoms

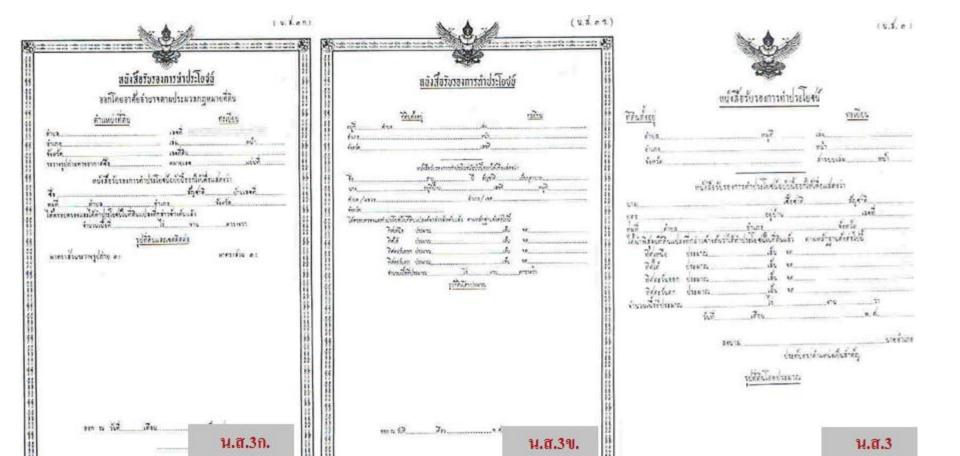


Who Does Data Belong to?



I will look at Three Types

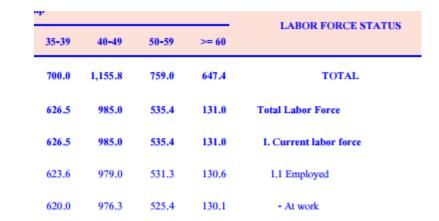
Public Statistics -National Statistics Office (NSO) of Thailand



History of the Thai Labor Force Survey



A Book of Statistics



CD, with Photos of Pages 😊



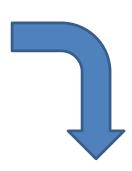
โตรมาสที่ 3/2558 : Quarter 3/2015										หน่วยเป็นพัน : In Thousan
	รวม				กลุ่มอายุ A	ge Group				LABOR FORCE STATUS
สถานภาพแรงงาน	Total	15-19	20-24	25-29	30-34	35-39	40-49	50-59	>= 60	LABOR FORCE STATUS
ประชากรอายุ 15 ปีขึ้นไป	55,290.6	4,684.6	4,946.2	4,756.5	4,862.0	5,295.7	10,915.3	9,357.7	10,472.6	POPULATION 15 YEARS AND OVER
กำลังแรงงานรวม	38,766.5	898.8	3,242.1	4,139.2	4,364.0	4,757.8	9,735.1	7,700.7	3,928.9	Total Labor Force
1. กำลังแรงงานปัจจุบัน	38,686.4	893.4	3,234.3	4,134.9	4,361.7	4,749.9	9,719.6	7,679.7	3,912.9	1. Current labor force
1.1 ผู้มีงานทำ	38,330.4	844.2	3,082.2	4,062.7	4,339.5	4,730.0	9,696.9	7,666.1	3,908.8	1.1 Employed
- ท้างาน	38,123.9	841.5	3,075.1	4,047.4	4,325.0	4,709.0	9,641.4	7,617.4	3,867.3	- At work

PDF, with selectable text

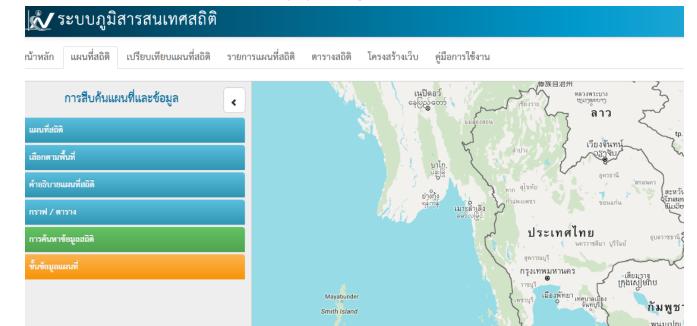
FILE HISTORY OF THE PAT Hai Labor Force Survey



Format Painter



Mapping of Selected Statistics



Data Repository with Filtering Capability



View data catalog »

Next step for big budget projects...

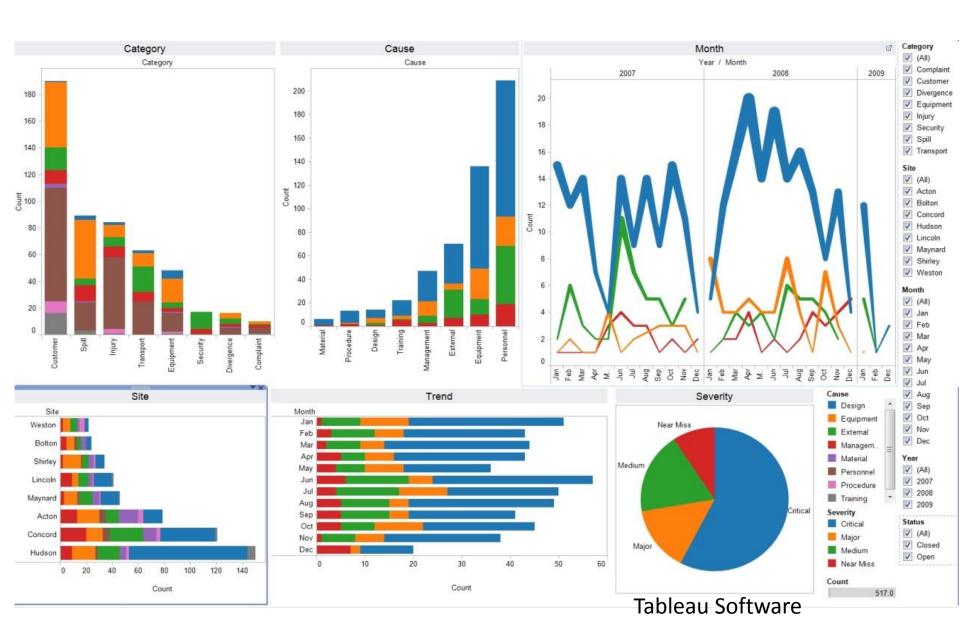
Desirable next step:

Query the Original Data online

World Bank, National Statistics Office select the variables to summarize for us

By linking to statistical software (R, Stata, SPSS) we could study new relationships without confidentiality issues, and without releasing original data.

?? Business Intelligence Platform ??



Thailand – 8 Principles of Good Practices in Official Statistics

- 1) Relevance, Impartiality and Equal Access
- 2) Professionalism (collecting in fair way)
- 3) Accountability (professional statistics)
- 4) Prevention of Misuse
- 5) Cost-Effectiveness
- 6) Confidentiality
- 7) Legislation (Transparency)
- 8) National Co-Ordination

Text from Principle 1

 Official statistics provide an indispensable element in the information system of a society, serving the government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honor citizen's entitlements to public information.

How can this compromise be reached?

- Data is collected in surveys, often at great expense and at a scale of the tens of thousands.
- To release the raw data creates a risk of compromising principles 4 (misuse) and 6 (privacy) above.
- Data has already been anonymized but the potential to unravel it exists.
- (NSO compromise Release to province level but not lower)
- Data depreciates rapidly.
- Data can be useful for many people in widely differing and unpredictable ways.

Solution may be to Release data only as aggregate statistics, but provide flexibility about what is measured. – Thus Data Visualization could allow for better access

 No statistic may be released if it is based on the aggregation of less than 20 (50) items

Average Income for Districts of Chonburi Province					
District	Income	Sampled	If N=20	If N=50	
1	23,049	129	23,049	23,049	
2	16,499	34	16,499	N/A	
3	16,644	9	N/A	N/A	
4	23,033	134	23,033	23,033	
5	16,020	57	16,020	16,020	
6	22,282	49	22,282	N/A	
7	18,686	42	18,686	N/A	

หน่วยราชการอิสระ

สำนักราชเลขาธิการ	http://www.ohmpps.go.th/	
สำนักพระราชวัง	http://www.palaces.thai.net/	
สำนักงานพระพุทธศาสนาแห่งชาติ	http://www.onab.go.th	
สำนักงานคณะกรรมการพิเศษเพื่อประสานงาน โครงการอันเนื่องมาจากพระราชดำริ	http://www.rdpb.go.th/	
สำนักงานคณะกรรมการวิจัยแห่งชาติ	http://www.nrct.go.th/	
ราชบัณฑิตยสถาน	http://www.royin.go.th/	
สำนักงานตำรวจแห่งชาติ	http://www.police.go.th/	
สำนักงานป้องกันและปราบปรามการฟอกเงิน	http://www.amlo.go.th/	
สำนักงานอัยการสูงสุด	http://www.ago.go.th/	



สำนักนายกรัฐมนตรี : thaigov.go.th

	http://www.opm.go.th/		
กรมประชาสัมพันธ์	http://www.prd.go.th/		
สำนักงานคณะกรรมการคุ้มครองผู้บริโภค	http://www.ocpb.go.th/		
	http://spm.thaigov.g		
สำนักเลขาธิส สคณะรัฐม หรื	http://www.cabinet.tra.pv.go.tl		
สำนักข่าวกราแห่งชาติ	h p:// ww.nia.go.th		
สำนักงบประเทย	neep:// www.bb.go.th/		
	http://www.nsc.go.th		
สำนักงานคณะกรรมีการกฤษฎีกา	http://www.krisdika.go.th/		
สำนักงานคณะกรรมการข้าราชการพลเรือน	http://www.ocsc.go.th/		
สำนักงานคณะกรรรมการพัฒนาการเศรษฐกิจ และสังคมแห่ง ชาติ	http://www.nesdb.go.th/		
สำนักงานคณะกรรมการพัฒนาระบบราชการ (ก.พ.ร.)	http://www.opdc.go.th/		



กระทรวงกลาโหม : mod.go.th

	h
สำนักงานปลัดกระทรวงกลาโหม	http://www.mod.go.th/opsd/
กรมราชองครักษ์ กระทรวงกลาโหม	http://www.radc.go.th/
องค์การสงเคราะห์ทหารผ่านศึก	http://www1.mod.go.th/opsd/wvoweb/
กองบัญชาการทหารสูงสุด	http://www.schq.mi.th/
กองทัพบก	http://www.rta.mi.th/
กองทัพเรือ	http://www.navy.mi.th/
กองทัพอากาศ	http://www.rtaf.mi.th/
องค์การแบตเตอรี่	http://powerbattery.thaigov.net/
องค์การฟอกหนัง	http://www.tto.or.th/



กระทรวงการคลัง : mof.go.th

สำนักงานปลัดกระทรวงการคลัง	http://palad.mof.go.th/
กรมธนารักษ์	http://www.treasury.go.th/
กรมบัญชีกลาง	http://www.cgd.go.th/
กรมศุลกากร	http://www.customs.go.th/
กรมสรรพสามิต	http://www.exd.mof.go.th/
กรมสรรพากร	http://www.rd.go.th/
สำนักงานคณะกรรมการนโยบายรัฐวิสาหกิจ	http://www.privatisation.go.th/
สำนักงานบริหารหนี้สาธารณะ	http://www.pdmo.mof.go.th/
สำนักงานเศสษาที่การกลัง	http://www.fpo.go.th/



กระหรวงศึกษาธิการ : moe.go.th

ข่าวสำนักงานรัฐมนตรี	http://www.moe.go.th/websm/news-online.htm
สำนักงานปลัดกระทรวงศึกษาธิการ	http://www.moe.go.th/OPS_Page/
สำนักงานเฉขาธิการสภาการศึกษา	http://www.onec.go.th/
สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน	http://www.obec.go.th/
สำนักงานคณะกรรมการการอุดมศึกษา	http://www.mua.go.th/
ส่วกัดสวนอดเหตุธรมดาธุดาธุกาศักดิ์ดหา	http://www.vac.go.th/



กระทรวงสาธารณสุข : moph.go.th

สำนักงานปลัดกระทรวงสาธารณสุข	
กรมการแพทย์	http://www.dms.moph.go.th/
กรมควบคุมโรค	http://www-ddc.moph.go.th/
กรมพัฒนาการแพทย์แผนไทย และการแพทย์ทางเลือก	http://203.157.19.1/dtam/
กรมวิทยาศาสตร์การแพทย์	http://www.dmsc.moph.go.th/
กรมสนับสนุนบริการสุขภาพ	http://hss.moph.go.th/
กรมสุขภาพจิต	http://www.dmh.moph.go.th/
กรมอนามัย	http://www.anamai.moph.go.th/
สำนักงานคณะกรรมการอาหารและยา	http://www.fda.moph.go.th/



กระทรวงการต่างประเทศ : mfa.go.th

สำนักงานปลัดกระทรวงการต่างประเทศ	http://www.mfa.go.th/web/1724.php
กรมการกงสุล	http://www.consular.go.th/
กรมพิธีการทูต	http://203.150.20.1/datapr/adt14.htm
กรมยุโรป	http://203.150.20.1/datapr/adt15.htm
กรมวิเทศสหการ	http://www.dtec.thaigov.net/
กรมเศรษฐกิจระหว่างประเทศ	http://203.150.20.1/datapr/adt16.htm
กร สนธิสัญญาแ ะกฎ 🕒 เาย	http://203.150.20.1/datapr@dt17.htm
กรสารมีเทศ	tn://203.150.20.1 atapr/adt18.htm
กร เงค์ก ระหว บประ าศ	tp://203.15 20.1 atapr, dt1 ftm
กรอเมริก ละ ไซิฟิได้	tp://200322020.1 atapr, dt20
กระวาเซีย	tp;// (3.15 20.1 atapr) dt21.htm
กรมเอเซยตะวันออก	http://205.150.20.1/datapr/adt22.nen
กรมเอเชียใต้ ตะวันออกกลางและแอฟริกา	http://203.150.20.1/datapr/adt23.htm





กระทรวงการพัฒนาสังคมและความมั่นคงของมนุษย์ : m-society.go.th

สำนักงานปลัดกระทรวงการพัฒนาสังคม	-
กรมพัฒนาสังคมและสวัสดิการ	http://www.dsdw.go.th/
สำนักงานกิจการสตรีและสถาบันครอบครัว	http://www.thaiwomen.net/
สำนักงานส่งเสริมสวัสดิภาพ และพิทักษ์เด็ก เยาวชน ผู้ด้อย โอกาส คนพิการ และผู้สูงอายุ	http://www.opp.go.th/



พรวงเกษตรและสหกรณ์ : moac.go.t

สำนักงานปลัดกระทรวงเกษตรและสหกรณ์	-
กรมชลประทาน	http://www.rid.go.th/
กรมตรวจบัญชีสหกรณ์	http://www.cad.go.th/
กรมประมง	http://www.fisheries.go.th/
กรมปศุลัตว์	http://www.dld.go.th/
กรมป่าไม้	http://www.forest.go.th/
กรมพัฒนาที่ดิน	http://www.ldd.go.th/
กรมวิชาการเกษตร	http://www.doa.go.th/
กรมส่งเสริมการเกษตร	http://www.doae.go.th/
กรมส่งเสริมสหกรณ์	http://www.cpd.go.th/
สำนักงานการปฏิรูปที่ดินเพื่อเกษตรกรรม	http://www.alro.go.th/
สำนักงานมาตรฐานสินค้าเกษตร และอาหารแห่งชาติ	http://www.napsi.or.th/
สำนักงานเศรษฐกิจการเกษตร	http://www.oae.go.th/



กระทรวงคมนาคม : mot.go.th

สำนักงานปลัดกระทรวงคมนาคม	-
กรมการขนส่งทางน้ำและพาณิชยนาวิ	http://www.md.mot.go.th/
กรมขนส่งทางบก	http://www.dlt.mot.go.th/
กรมการขนส่งทางอากาศ	http://www.aviation.go.th/
กรมทางหลวง	http://www.doh.mot.go.th/
กรมทางหลวงชนบท	http://www.dor.go.th/
สำนักงานนโยบายและแผนการขนส่งจราจร	http://www.otp.go.th/



กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม : mnre.go.th

ล่านักงานปลัดกระทรวงกระทรวง ทรัพยากรธรรมชาติ และสิ่ง แวดล้อม	-
กรมควบคุมมลพิษ	http://www.pcd.go.th/
กรมทรัพยากรทางทะเลและชายฝั่ง	http://www.monre.go.th/product_4.html
กรมทรัพยากรธรณี	http://www.dmr.go.th/
กรมทรัพยากรน้ำ	http://www.dwr.go.th/
กรมทรัพยากรน้ำบาดาล	http://www.dgr.go.th/
	http://www.monre.go.th/product_9.html
กรมส่งเสริมคุณภาพสิ่งแวดล้อม	http://www.deqp.go.th/
สำนักงานนโยบายและแผนทรัพยากรธรรมชาติ และสิ่ง แวดล้อม	http://www.oepp.go.th/



ะทรวงเทคโนโลยีสารสนเทศและการสื่อสาร : mict.go.th

สำนักงานปลัด กระทรวงเทคโนโลยีสารสนเทศ และการสื่อสาร	-
กรมไปรษณีย์โทรเลข	http://www.ptd.go.th/
กรมอุตุนิยมวิทยา	http://www.tmd.go.th/
สำนักงานสถิติแห่งชาติ	http://www.nso.go.th/



_

กระทรวงพลังงาน : energy.go.th

	สานกงานปลดกระทรวงพลงงาน	-
	กรมเชื้อเพลิงธรรมชาติ	http://www.dmf.go.th/
	กรมธุรกิจพลังงาน	http://www.doeb.go.th/
	กรมพัฒนาพลังงานทดแทน และอนุรักษ์พลัง <u>งา</u> น	http://www.dedp.go.th/
_	ສ່ານັກນໂບນາຍແລະແຜນຫລັນງນ	http://www.eppo.go.th/
	MARTI	MANTC
W		
٠,		
IV.		
	• nl	

กระหรวงพาณิชย์ : moc.go

สำนักงานปลัดกระทรวงพาณิชย์	http://www.moc.go.th/thai/opsmoc/ops1/index.html
กรมการค้าต่างประเทศ	http://www.dft.moc.go.th/
กรมการค้าภายใน	http://www.dit.go.th/
กรมการประกันภัย	http://www.moc.go.th/thai/doi/
กรมเจรจาการค้าระหว่างประเทศ	http://dtn.moc.go.th/
กรมทรัพย์สินทางปัญญา	http://www.ipthailand.org/
กรมพัฒนาธุรกิจการค้า	http://www.thairegistration.com/
กรมส่งเสริมการส่งออก	http://www.depthai.go.th/



ระทรวงมหาดไทย : moi.go.th

สำนักงานปลัดกระทรวงมหาดไทย	-
กรมการปกครอง	http://www.dopa.go.th/
กรมการพัฒนาชุมชน	http://www.cdd.moi.go.th/
กรมที่ดิน	http://www.dol.go.th/
กรมป้องกันและบรรเทาสาธารณภัย	http://www.dpm.moi.go.th/
กรมโยธาธิการและผังเมือง	http://www.dpt.go.th/
กรมส่งเสริมการปกครองท้องถิ่น	http://www.thailocalgov.net/



กระหรวงยุติธรรม : moj.go.th

สำนักงานปลัดกระทรวงยุติธรรม	-	
กรมคุมประพฤติ	http://www.probation.go.th/	
กรมคุ้มครองสิทธิและเสริภาพ	http://www.rlpd.moj.go.th/	
กรมบังคับคดี	http://www.led.go.th/	
กรมพินิจและคุ้มครองเด็กและเยาวชน	http://www.djop.moj.go.th/	
กรมราชทัณฑ์	http://www.correct.go.th/	
กรมสอบสวนคดีพิเศษ	http://www.sid.moj.go.th/	
สำนักงานกิจการยุติธรรม	http://www.oja.moj.go.th/	
สถาบันนิติวิทยาศาสตร์	http://www.cifs.moj.go.th/	
สำนักงานคณะกรรมการป้องกัน และปราบปรามยาเสพติด	http://www.oncb.go.th/	



กระทรวงแรงงาน : mol.go.th

สำนักงานปลัดกระทรวงแรงงาน	-
กรมการจัดหางาน	http://www.doe.go.th/
กรมพัฒนาฝีมือแรงงาน	http://www.dsd.go.th/
กรมสวัสดิการและคุ้มครองแรงงาน	http://www.labour.go.th/
สำนักงานประกันสังคม	http://www.sso.go.th/



าระทรวงวัฒนธรรม : m-culture.go.th

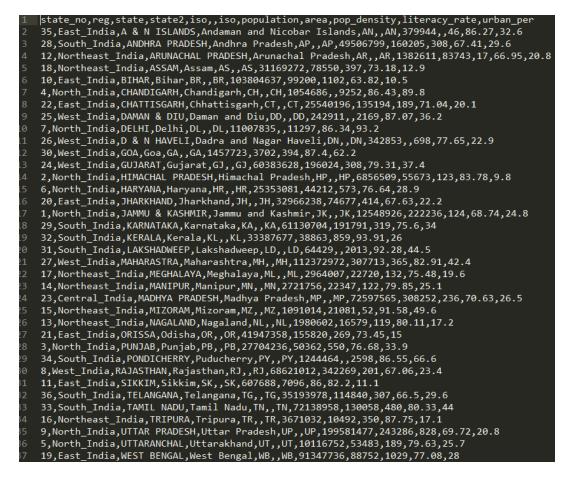
สำนักงานปลัดกระทรวงวัฒนธรรม	http://www.culture.go.th/ops/
กรมการศาสนา	http://www.moe.go.th/webrad/
กรมศิลปากร	http://www.finearts.go.th/
สำนักงานคณะกรรมการวัฒนธรรมแห่งชาติ	http://www.culture.go.th/oncc/
สำนักงานศิลปวัฒนธรรมร่วมสมัย	http://www.culture.go.th/omac/

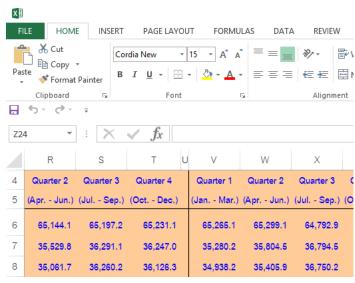
Smaller Ministries and Departments Desirable next step...Improved Access

- Depends on size and budget of agency.
- Same 8 principles apply, but for many of them the capacity and budgets are very low.
- All agencies are collecting data at least at the province level, and often in much more detail.
- Information would be of use and interesting to many in the general public.
- Very little of the data is publicly accessible unless requested by an academic or business.
- Even if available, tables of statistics are hard to read and interpret.

My Current Project

Providing a very simple visualization tool to link
 CSV data with a map or graphic, as a goodwill way to allow small ministries to share data





Lack of Capacity

- Everywhere I go, I get the same story We have lots of data – should be useful – don't really know how to use it.
- Government officers are overworked, and underpaid.
- They fly to many provinces to collect statistics and consult experts and farmers
- May be well trained, but not in computer.
- Solution: Provide a simple free solution to Data Visualization such as a map tied to an excel file.

Privacy Concerns

 No one can know about the yield on Somchai's corn field







• Solution: Use aggregation (20) and heat map

Data is not "clean" and we don't have time to do it.

- Data is provisional
- Data is incomplete
- Somchai's yield was 2.3-2.6 tons per hectare.
 - What number to use in a range?
- Daeng's yield is in kilos per rai
- **Solution:** Using some type of heat map means that many details can be ignored.
- Besides this, standardized forms will help

Data as a source of power

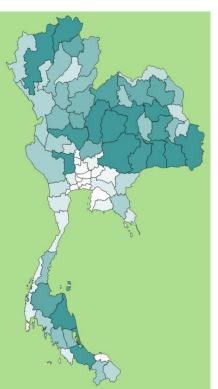
- Within the ministry, each person has a very narrow window of responsibility.
- Holding data on a particular topic may give that person power or importance.
- Making data public reduces their position

• **Solution:** Releasing data must be ministry wide and authority must come from the top.

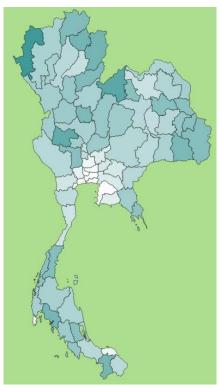
Agricultural Bank – Hero or Villain?

It looks like Agricultural Bank sets up offices where there are many farmers (picture 1) Once there, they make loans to both farmers and non-farmers, and also run the village fund. The result is a high rate of indebtedness.

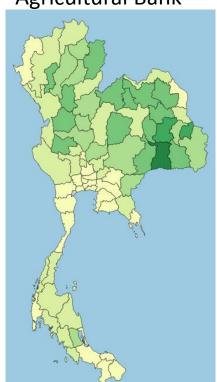
Number of Farmers in Province



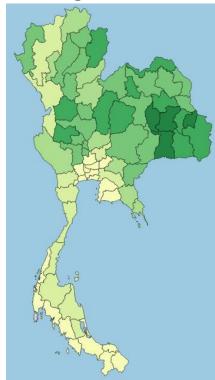
Percent of Population who are Farmers



Percent who owe money to Agricultural Bank

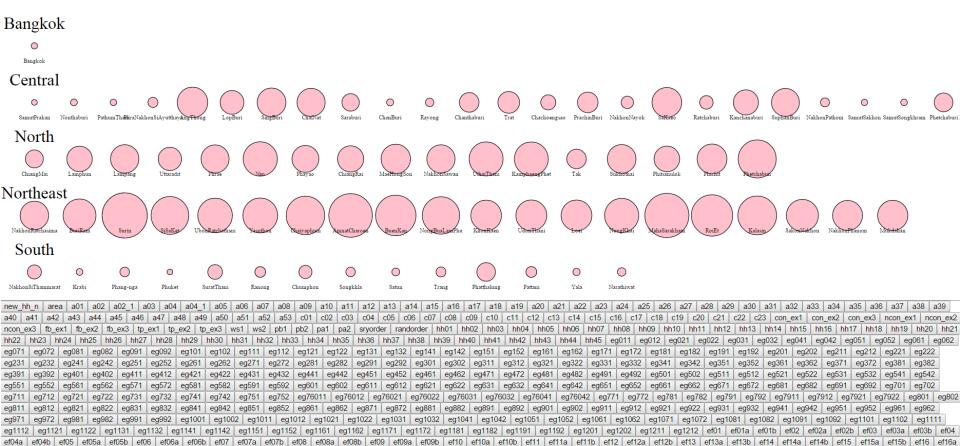


Percent who owe money to Village Fund



Village Fund

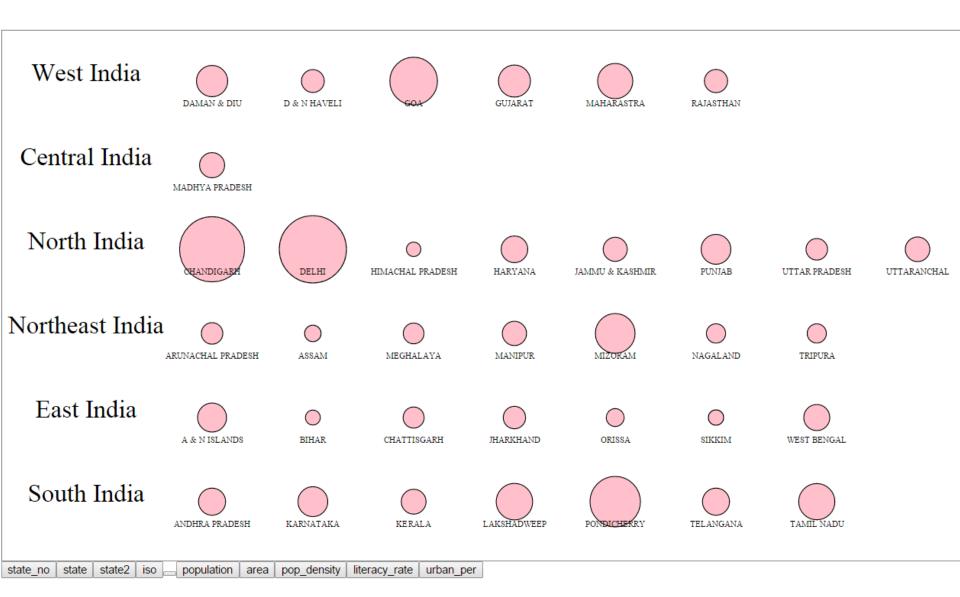
Is linked to the Agricultural Bank



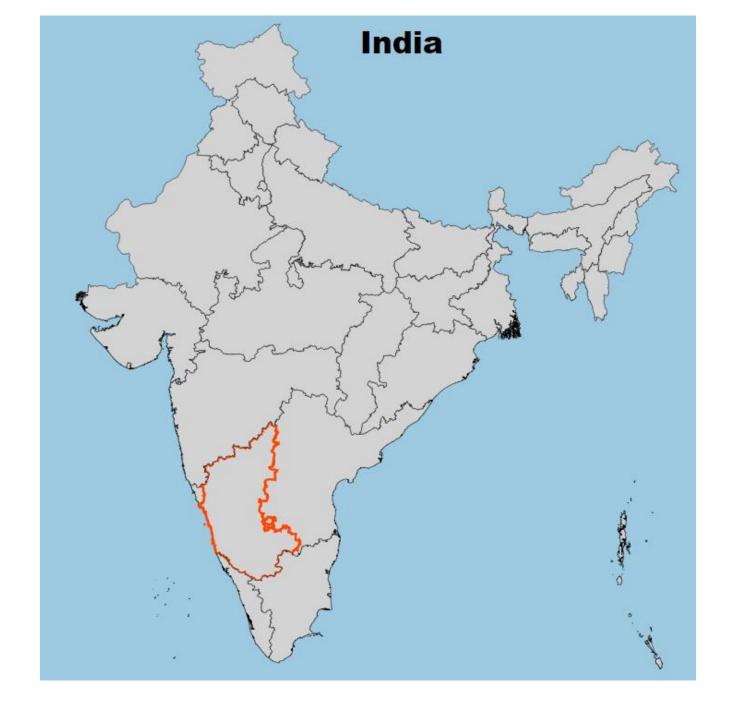
ia032 | ia033 | ia041 | ia042 | ia051 | ia052 | ia061 | ia062 | ia061 | ia062 | ia07 | ia08 | ia09 | ia10 | ia11 | ia12 | ia13 | ia14 | ia15 | ia16 | ia17 | ia18 | ia19 | ia20 | ia21 | ia22 | ia23 | ia24 | ia25 | ia26 | ia27



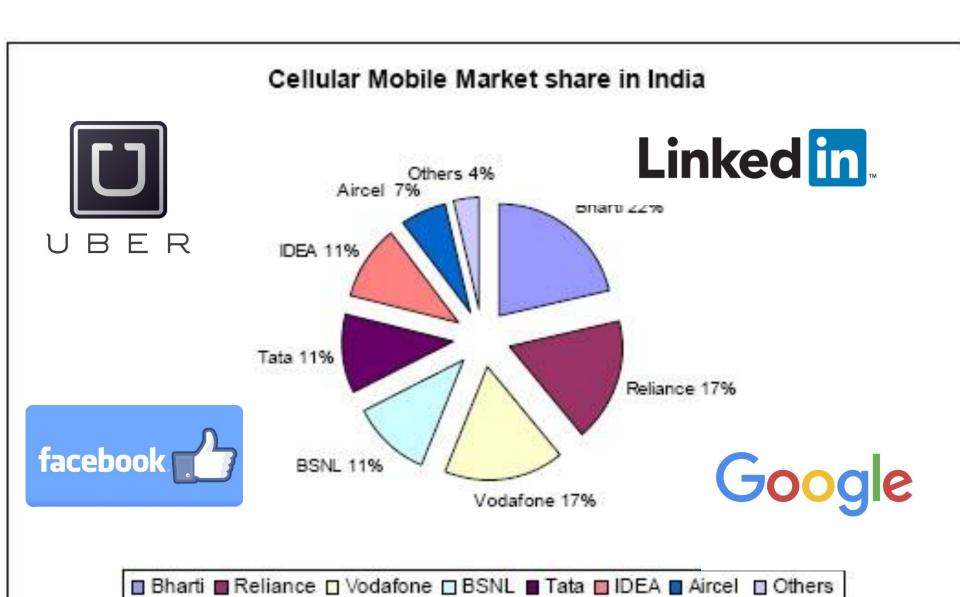
a0	2	a02	_1	a03	3 a0)4	a04	_1	a05		a06	a07	7 a	80	a()9	a1	0
38	9	a40	a ⁴	11	a42	a	43	a44	a4	5	a46	a4	7	a48	a	49	a	50
n_	ex2	2 c	on_e	ex3	nco	n_e	x1	nco	n_ex	2	nco	n_ex	3 1	fb_e	x1	fb	ex	2
) [hh	11	hh1	2	hh13	H	nh14	h	h15	hł	n16	hh1	7 1	hh18		hh1	9	hhí
3	hh	44	hh4	15	eg01	1	eg0	12	eg0	21	eg()22	eg(031	e	g03:	2	eg(
114	11	eg	142	eg	151	eg	152	e	g161	E	eg162	2 e	g171	l e	g1	72	eg	181



Button name	Group	State	Year	population	area	pop_density lit	eracy_rate urba	n_per
Description	East_India	A & N ISLANDS	2013	379944		46	86.27	32.6
	South_India	ANDHRA PRADESH	2013	49506799	160205	308	67.41	29.6
	Northeast_India	ARUNACHAL PRADESH	2013	1382611	83743	17	66.95	20.8
	Northeast_India	ASSAM	2013	31169272	78550	397	73.18	12.9
	East_India	BIHAR	2013	103804637	99200	1102	63.82	10.5
	North_India	CHANDIGARH	2013	1054686		9252	86.43	89.8
	East_India	CHATTISGARH	2013	25540196	135194	189	71.04	20.1
	West_India	DAMAN & DIU	2013	242911		2169	87.07	36.2
	North_India	DELHI	2013	11007835		11297	86.34	93.2
	West_India	D & N HAVELI	2013	342853		698	77.65	22.9
	West_India	GOA	2013	1457723	3702	394	87.4	62.2
	West_India	GUJARAT	2013	60383628	196024	308	79.31	37.4
	North_India	HIMACHAL PRADESH	2013	6856509	55673	123	83.78	9.8
	North_India	HARYANA	2013	25353081	44212	573	76.64	28.9
	East_India	JHARKHAND	2013	32966238	74677	414	67.63	22.2
	North_India	JAMMU & KASHMIR	2013	12548926	222236	124	68.74	24.8
	South_India	KARNATAKA	2013	61130704	191791	319	75.6	34
	South_India	KERALA	2013	33387677	38863	859	93.91	26
	South_India	LAKSHADWEEP	2013	64429		2013	92.28	44.5
	West_India	MAHARASTRA	2013	112372972	307713	365	82.91	42.4
	Northeast_India	MEGHALAYA		2964007	22720	132	75.48	19.6



Lastly, Private Statistics -



Fair Usage – Basis for EU Privacy Law

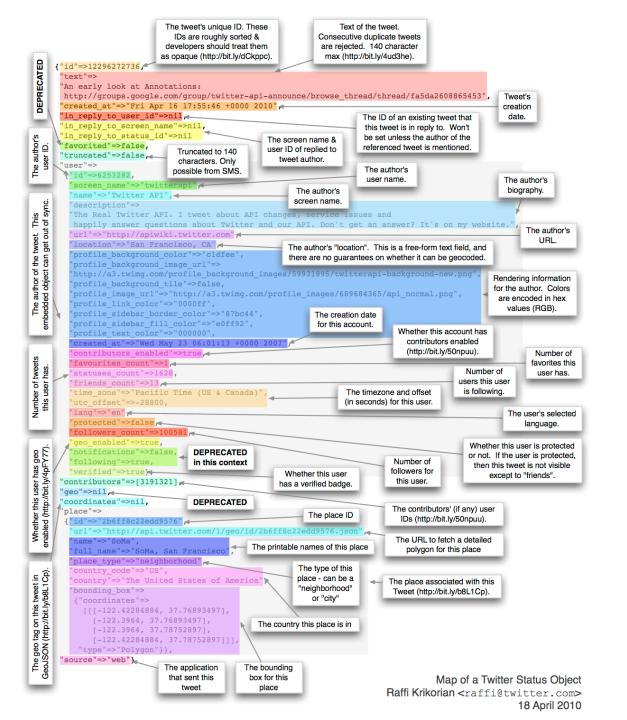
NOT TRUE

- For all data collected there should be a stated purpose.
- Information collected by an individual cannot be disclosed to other organizations or individuals unless specifically authorized by law or by <u>consent</u> of the individual <u>Okay</u> - <u>Terms of Usage</u>
- Records kept on an individual should be accurate and up to date impractical
- There should be mechanisms for individuals to review data about them, to ensure accuracy. This may include periodic reporting
- Data should be deleted when it is no longer needed for the stated purpose
 NOT TRUE
- Transmission of personal information to locations where "equivalent" personal data protection cannot be assured is prohibited

 Maybe Untrue
- Some data is too sensitive to be collected, unless there are extreme circumstances (e.g., sexual orientation, religion) Impractical, but Best Effort

In every transaction there is a Buyer and a Seller

- Both should have access to data from the transaction.
- Data that is not relevant to the transaction should not be collected
- Data that is not needed for future reference should not be kept.
- BUT, WHAT ABOUT SERVICES THAT DO NOT CHARGE A FEE, AND SELL THE DATA?



 A Twitter text is limited to 140 characters, but the metadata that accompanies each twitter text is about 1500 characters!

Compromise – Shared Data?

- Data can be collected if it can be shown to be useful to both parties in the transaction
- (E.G. as Google often does)

• **Data as a club good**, in which customers are part of the club.

Should Data be Treated as a Monopoly?

Information Monopoly Act

Given that many technology products are now necessary to lead a normal human life, and given that these technology products are provided by private companies, which are by their nature necessarily natural monopolies, there is a need for the government to regulate the collection and distribution of data so that it is non-abusive, fair for the company and the public on which data is collected, and available in anonymous or aggregated form for a fee to third parties and the government for purposes of the public good.

In line with the above, the following principles should be guaranteed

- 1) Some return for those who collect data in line with cost of collection
- 2) Some ability for others to use data prices subject to approval by the government
- 3) Some ability for governments to make use of data for public good projects
- 4) Ability for companies to preserve original data, that can be licensed.
- 5) Ability to reserve data that is specific to the company for competitive reasons.

Indian Statistical Institute, Bangalore 2016