



รศ.ดร. อภิญา หิรัญวงษ์
ภาควิชาสถิติ คณะวิทยาศาสตร์ วิทยาเขตบางเขน มก.
E-mail - fsciaph@ku.ac.th

1/2557

1 ความสำคัญของสถิติ

1 'Daily life decision making involves statistics'
'การตัดสินใจในชีวิตประจำวันเกี่ยวข้องกับสถิติ'

'Statistics is around'

สถิติในชีวิตประจำวันมีมากมายไม่จำกัดขอบเขตว่าอยู่ในเรื่องใดเรื่องหนึ่ง และการใช้สถิติอย่างไม่รู้ตัวอยู่เป็นประจำ



การอนุมานสถิติในชีวิตประจำวัน

- 1 ความสำคัญของสถิติ
- 2 ข้อมูลสถิติและการเก็บรวบรวมข้อมูล
- 3 การประมวลผลข้อมูลเชิงพรรณนา



ตะลึง! การศึกษาไทย อันดับ 7 ของอาเซียน

(ผลสำรวจปี 2014-2015)

ความสามารถในการแข่งขันของไทย
อยู่ในอันดับที่ 31 ของโลก

คุณภาพการศึกษาขั้นพื้นฐานของไทย
ตกลงมาอยู่ อันดับที่ 7 ของอาเซียน และ อันดับที่ 86 ของโลก

คุณภาพการศึกษาระดับอุดมศึกษาของไทย
อยู่ใน อันดับที่ 8 ของอาเซียน และ อันดับที่ 78 ของโลก

คุณภาพ คนิต-วิทย์ ของไทย
อยู่ใน อันดับที่ 5 ของอาเซียน และ อันดับที่ 80 ของโลก



ความหมายของสถิติ

1. ตัวเลข
กลุ่มตัวเลขหรือตัวเลข ข้อมูลสถิติ
2. ศาสตร์หรือวิชา
การรวบรวมข้อมูล การจัดระเบียบข้อมูล การนำเสนอข้อมูล การวิเคราะห์ การแปลความหมายและตีความ การวิเคราะห์เชิงสถิติระดับสูง



อัตราแลกเปลี่ยน		ราคาเชื้อเพลิง	
ดอลลาร์สหรัฐ	32.48 บาท	เบนซิน 95	44.86 บาท
เยน ญี่ปุ่น(100)	30.75 บาท	น้ำมันดีเซล	29.99 บาท
ยูโร	42.02 บาท	แก๊สโซฮอล์ 95	37.80 บาท
ปอนด์ อังกฤษ	53.04 บาท	แก๊สโซฮอล์ 91	35.78 บาท
หยวน จีน	5.40 บาท	แก๊สโซฮอล์ E20	33.98 บาท
ดอลลาร์ฮ่องกง	4.24 บาท	LPG	280-290/15 kg บาท
ดอลลาร์ออสเตรเลีย	29.82 บาท	NGV	10.50 บาท
ดอลลาร์ สิงคโปร์	25.73 บาท	NVME	92.83 ดอลลาร์สหรัฐ
ราคาทองคำ	รับซื้อ	ขายออก	
ทองคำแท่ง	18,800.00 บาท	18,900.00 บาท	
ทองรูปพรรณ	18,525.52 บาท	19,300.00 บาท	

http://www.manager.co.th/home ผู้จัดการออนไลน์ ฉบับวันที่ 13 กันยายน 2557

ตัวอย่าง ประโยชน์ของสถิติ

• **การบริหารงานและพัฒนาประเทศ**

ข้อมูลสถิติมีความสำคัญและจำเป็น โดยเป็นเครื่องมือ สำหรับผู้บริหาร ใช้เป็นแนวทางประกอบการตัดสินใจในการจัดทำแผนงาน กำหนดนโยบายหรือแก้ไขปัญหา

- โอกาสที่จะมีฝนตกในวันหนึ่งๆ
- การทดสอบประสิทธิภาพของยาโรก
- การควบคุมคุณภาพของสินค้า
- การสำรวจความคิดเห็นหรือโพล

วารสาร วิชาการ "กรนอน" นักวิจัยของสองชาติได้ ร่วมกันศึกษาความ เกี่ยวพันของการนอนกับอัตราการเสียชีวิต ด้วยการทบทวนรายงาน การศึกษาเกี่ยวกับการนอน 16 เรื่อง ซึ่งเกี่ยวข้องกับผู้คนมากถึง 1,500,000 คน ได้ข้อสรุปว่า "การเสียชีวิตก่อนวัยอันควรทั้งหมด ล้วน แต่เกี่ยวข้องกับการนอนน้อยไปหรือนอนมากเกินไป มากน้อยเกินกว่าเวลา นอนที่เหมาะสมที่ควร คินละระหว่าง 6-8 ชม. ทั้งสิ้น" คณะ นักวิจัย เชื่อว่า ขณะที่การนอนน้อยอาจจะเป็นสาเหตุของสุขภาพเสื่อมโดยตรง ซึ่งนำไปสู่การมีอายุสั้นที่สุดในที่สุด แต่การนอนนานก็อาจเป็นเพียงเครื่อง บอกให้รู้ว่า สุขภาพไม่ดี.

ข้อมูลข่าวโดยหนังสือพิมพ์ไทยรัฐออนไลน์ วันที่ 8 พฤษภาคม 2553

2 สถิติเป็นเครื่องมือในการวิจัยทุกสาขาวิชา

สำหรับวิเคราะห์ข้อมูลเพื่อให้ได้ผลสรุปตาม วัตถุประสงค์ที่กำหนด

การใช้สถิติในการวิจัยทำให้ผลการวิจัยน่าเชื่อถือและ ยอมรับในวงวิชาการ

ตัวอย่างงานวิจัย ที่ใช้สถิติเป็นหลักฐานซึ่งเป็นข้อมูลเชิง ประจักษ์และใช้วิธีการทางสถิติในการสรุปผลการวิจัย

ดื่มชามากอาจ "ฆ่า" ไม้ออกเพราะแก่เร็ว

โดย ASTVผู้จัดการออนไลน์ 6 พฤษภาคม 2553 00:36 น. [Tweet](#) (1) [G+](#) (0) [Share](#) (4) [Facebook](#) 4 [Google+](#) 0 [Twitter](#) 1



การศึกษานักวิจัยสหรัฐฯ พบว่าการที่ผ่านกระบวนการและโรคที่เสริมฟอสเฟตเพื่อ "ความขุ่น" นั้น ได้เร่งสัญญาณ "ความ หนืด" ในหลอดเลือด อาทิ โรคไต ภาวะแคลเซียมเกาะหนึ่เส้นเลือดหัวใจ อีกทั้งยังเร่งความเสื่อมต่อกล้ามเนื้อและผิวหนังอย่างรุนแรง

"คนเราต้องกินอาหารเพื่อสุขภาพและการรักษาความสมดุลของฟอสเฟตในอาหาร อาจเป็น สิ่งสำคัญต่อการใช้สุขภาพดีและมีชีวิตยืนยาว หลีกเลียงฟอสเฟตที่เป็นพิษแล้วเป็นพิษกับชีวิตที่มี สุขภาพดี" ไซมอนด์ระบบศาสตร์ ดร.พ.ชอว์ท ราซเซค (Shawkat Razaque) จากภาควิชาการแพทย์ การติดเชื้อและภูมิคุ้มกัน (Department of Medicine, Infection and Immunity) จากมหาวิทยาลัยฮาร์วาร์ด (Harvard School of Dental Medicine) สหรัฐฯ

ราซเซคและคณะได้ศึกษาผลกระทบจากการได้รับฟอสเฟตปริมาณสูงในหนู 3 กลุ่ม กลุ่มแรกเป็นหนูที่ไม่มีฟัน "โคโธ" (kotho) ซึ่งเป็น สัตว์ชนิดหนึ่งซึ่งฟันเป็นพิษต่อร่างกาย ซึ่งหนูกลุ่มนี้มีชีวิตอยู่ได้ 8-15 สัปดาห์ กลุ่มต่อมาเป็นหนูที่ขาดฟันโคโธและฟัน Nap2a ขบขวาง ภายหลังฟอสเฟตในระดับต่ำ และหนูกลุ่มนี้มีอายุได้ 20 สัปดาห์ ส่วนกลุ่มสุดท้ายเป็นหนูที่มีฟันเช่นเดียวกับหนูกลุ่มที่ 2 แต่กลับได้รับ อาหารที่ฟอสเฟตสูง ซึ่งพบว่าหนูกลุ่มสุดท้ายตายในเวลา 15 สัปดาห์

งานวิจัยชี้ว่าฟอสเฟตมีความเป็นพิษต่อหนูและอาจมีผลกระทบต่อสัตว์เลี้ยงลูกด้วยนมชนิดอื่น ซึ่งรวมถึงมนุษย์ด้วย

"ไขมันคอเลสเตอรอลเป็นตัวบ่งชี้ของโรคหัวใจและหลอดเลือด แต่มีฟอสเฟตเป็นพิษโดยธรรมชาติ งานวิจัยชี้ว่าความสมดุล ของฟอสเฟตในร่างกายเรานั้นมีอิทธิพลต่อกระบวนการแก่ ดังนั้นอย่าให้ฟันแก่มีไขมัน" นพ.เจรัลด์ รัสส์แมน (Gerald Weissmann) บรรณธิการวารสารฟาสเซบ (FASEB) ซึ่งเผยแพร่ผลงานวิจัยนี้ได้ความเห็น

ทั้งนี้ งานวิจัยของราชเขตพระมหากษัตริย์ฟอสเฟตสูงในอาหารผ่านกระบวนการและเครื่องดื่มช็อคโกแลต รวมถึงน้ำอัดลมด้วย

ตัวอย่าง 1 ศึกษาวิจัยพบ ความสำคัญของการนอน นอนไม่ถึง 6 ชม. อายุสั้น

การนอนน้อยอาจจะเป็นสาเหตุของสุขภาพเสื่อมโดยตรง ซึ่งนำไปสู่การ มีอายุสั้นที่สุดในที่สุด นักวิจัยทางการแพทย์ทั้งเมือง มะกะโรนีและน้ำชา กล่าวเตือนคนที่อดหลับอดนอนว่า ผู้ที่นอนหลับไม่ ถึงคินละ 6 ชม. มี หวังจะได้ดีแต่สมอเร็ก่อนวัยอันควร คนนอนน้อยเช่นนั้นอยู่เป็นประจำ มีหวังอายุสั้น จะเสียชีวิตลงภายในช่วงเวลา 25 ปีข้างหน้า มากกว่าผู้ที่ นอนนานตามเวลาที่เหมาะสม ระหว่าง 6-8 ชม. กว่กันถึงร้อยละ 12 แต่ก็น่าแปลกใจเช่นกันที่พวกเขาอ้างพบว่าผู้ที่ นอนนานเกิน 9 ชม. ก็เสี่ยงกับอายุสั้นด้วยเช่นกัน

2 ข้อมูลและการเก็บรวบรวมข้อมูล

ข้อมูล (data) คือข้อเท็จจริงหรือรายละเอียดเกี่ยวกับค่าที่ต้องการ สนใจศึกษา ตัวเลข ตัวอักษร หรือ สัญลักษณ์

ได้จากกรนับ การวัด การสังเกตหรือการสัมภาษณ์

ข้อมูลสถิติ statistical data

- อุณหภูมิสูงสุดในแต่ละวันของประเทศไทย ตั้งแต่ 2530-2552
- ปริมาณน้ำฝนในแต่ละจังหวัดของประเทศไทย ตั้งแต่ 2540-2552
- จำนวนนักท่องเที่ยวชาวต่างประเทศที่เข้ามาเที่ยวในประเทศไทย ตั้งแต่ปี 2545 ถึง 2551
- น้ำหนักและส่วนสูงของนักเรียนชั้นประถมปีที่ 1 ห้อง 1

01999211 คณิตศาสตร์และคอมพิวเตอร์ใน
ชีวิตประจำวัน:รศ.ดร.อภิญา หิรัญวงษ์

ตัวอย่างข้อมูลจากในเว็บไซต์

Nationality	2013		2012		%Δ 2013/2012
	Number	%Share	Number	%Share	
East Asia	1,136,141	50.89	946,302	47.80	20.06
ASEAN	490,280	21.88	402,048	23.19	6.11
Brunei	1,163	0.05	678	0.03	101.21
Cambodia	25,823	1.14	30,576	1.53	-16.20
Indonesia	42,829	1.91	29,557	1.48	44.90
Laos	59,991	2.68	71,701	3.60	-16.33

2. ข้อมูลเชิงคุณภาพ qualitative data เป็นข้อมูลที่ไม่สามารถวัดเป็นตัวเลขได้ แต่แสดงข้อเท็จจริงตามคุณสมบัติ

คณะ – เกษตร ประมง วิทยาศาสตร์ ...
เพศ - ชาย หญิง
วุฒิการศึกษา – ป.ตรี ป.โท ป.เอก
ชั้นปี – 1 2 3 4
ความพอใจต่อระบบการลงทะเบียนแบบออนไลน์ของมหาวิทยาลัย – พอใจ ไม่พอใจ

ประเภทของข้อมูล

จำแนกตามลักษณะ

1. ข้อมูลเชิงปริมาณหรือข้อมูลจากการวัด
Quantitative data, Measurement data, Numerical data
เป็นข้อมูลที่สามารวัดได้เป็นตัวเลขและมีหน่วยแน่นอน

น้ำหนัก ความสูง อายุการใช้งานของหลอดไฟ
ผลผลิตของข้าว ค่าธรรมเนียมต่างๆ ค่าใช้จ่าย

3. ข้อมูลจำแนกประเภท categorical data
นับจำนวนข้อมูลจำแนกตามประเภทชนิดหรือลักษณะ เป็นความถี่ของข้อมูลในแต่ละกลุ่ม

จำนวนนิสิต มก. จำแนกตามเพศ
จำนวนนิสิต มก. จำแนกตามคณะ
จำนวนคนจำแนกตามกลุ่มเลือด
จำนวนรถยนต์จำแนกตามยี่ห้อ

1. ข้อมูลต่อเนื่อง (continuous data) เป็นข้อมูลตัวเลขที่มีค่าติดต่อกันเป็นช่วง เช่น น้ำหนัก คะแนนสอบ

2. ข้อมูลไม่ต่อเนื่อง (discrete data) ได้จากการนับตามสิ่งที่สนใจ เช่น จำนวนคนในครอบครัว จำนวนเสียงโทรศัพท์ที่เข้ามาในสำนักงาน จำนวนบุตรในครอบครัว จำนวนครั้งที่นิสิตขาดเรียน

การเก็บรวบรวมข้อมูล

1. การสำรวจ

การสำมะโน (census) สำรวจทั้งหมด
การสุ่มตัวอย่าง (sampling survey) สำรวจบางส่วน

‘ข้อมูลจากการสำรวจ (Survey data)’

ตัวอย่าง

1. การสำรวจเกษตรกรในจังหวัดอุษุขา
2. สุ่มสอบถามประชาชน 1000 คน เกี่ยวกับจำนวนครั้งในการชมภาพยนตร์ในโรงภาพยนตร์ต่อเดือน รายได้ อายุ เพศ การศึกษา

เป็นการศึกษาจากการสังเกตและสอบถามจากสภาพจริง เพราะไม่มีการควบคุมปัจจัยที่อาจส่งผลต่อข้อมูลที่รวบรวม



19

อันดับ	ประเทศ/เขตปกครองพิเศษ	จำนวนประชากร	วันที่มีผล ครั้งล่าสุด	% ประชากรเทียบ กับประชากรโลก	แหล่งอ้างอิง
—	จีน	6,989,500,000		100%	US Census - International Programs Department ↗
1	ประเทศอินเดีย	1,349,020,000	23 มกราคม 2012	19.3%	Chinese Population clock ↗
2	สหรัฐอเมริกา	1,265,250,000	23 มกราคม 2012	17.24%	Indian Population clock ↗
3	บราซิล	314,150,000	23 มกราคม 2012	4.49%	Official USA Population clock ↗
4	รัสเซีย	231,369,500		3.31%	ประมาณการสหประชาชาติ
5	อินโดนีเซีย	196,630,000	23 มกราคม 2012	2.8%	Official Pakistan population clock ↗
6	ปากีสถาน	174,748,000	23 มกราคม 2012	2.5%	Official Pakistan population clock ↗
7	บังกลาเทศ	162,221,000		2.32%	ประมาณการสหประชาชาติ
8	อิหร่าน	154,729,000		2.21%	ประมาณการสหประชาชาติ
9	เม็กซิโก	141,821,000	23 มกราคม 2012	2.03%	Russian State Statistical Service ↗
10	ญี่ปุ่น	127,590,000	1 สิงหาคม 2009	1.83%	Official Japan Statistics Bureau estimate ↗
11	เกาหลีใต้	107,550,697	กลางปี 2009	1.54%	India Population Census ↗
12	เวียดนาม	92,222,660	กลางปี 2009	1.32%	National Statistics Office medium projection ↗
13	เยอรมนี	90,549,390	กรกฎาคม 2011	1.3%	DA Facebook about Vietnam ↗
14	ฝรั่งเศส	82,046,000	31 พฤษภาคม 2008	1.17%	Federal Statistical Office ↗
15	ไทย	79,221,000	ธันวาคม 2008	1.13%	Ethnic Central Statistical Agency ↗
16	แคนาดา	78,945,000	4 สิงหาคม 2010	1.1%	Official Egyptian population clock ↗
17	ออสเตรเลีย	74,196,000		1.06%	ประมาณการสหประชาชาติ
18	แอฟริกาใต้	71,517,100	31 ธันวาคม 2008	1.02%	Turkey Statistical Institute estimate ↗
19	สหราชอาณาจักร	67,354,820	31 ธันวาคม 2010	0.96%	สำนักงานสถิติแห่งชาติ กระทรวงมหาดไทย (กรมการปกครอง) และกรมการโยธาธิการ
20	อิตาลี	66,020,000		0.94%	ประมาณการสหประชาชาติ
21	เกาหลีใต้	65,073,482	1 มกราคม 2009	0.93%	Official INSEE estimate ↗

2. การทดลอง กำหนดสถานการณ์หรือการทดลอง
'ข้อมูลที่ได้จากการทดลอง (Experimental data)'

สุ่มตัวอย่างจากลูกค้า 400 คน แบ่งเป็น 4 กลุ่มๆละ 100 คน กำหนดให้ใช้แชมพูสระผม 4 ยี่ห้อ สอบถามความพอใจในการใช้แชมพูทั้ง 4 ยี่ห้อ

เป็นการศึกษาจากการทดลอง เพราะมีการควบคุมปัจจัยจำนวน 1 ปัจจัย คือ แชมพูที่ลูกค้าใช้กำหนด 4 ยี่ห้อ

ข้อมูลปฐมภูมิ (primary data)



20

ตัวเลขพื้นฐานทางเศรษฐกิจของประเทศอาเซียนทั้ง 10 ประเทศ

ประเทศสมาชิกอาเซียน	ประชากร (ล้านคน)	จีดีพี 2011 (ล้านดอลลาร์)	การขยายตัวของจีดีพี (ปี 2011)	จีดีพีต่อหัว (ดอลลาร์)	มูลค่าส่งออก (ล้านดอลลาร์)	ส่งออก เป็น% ของจีดีพี	ทุนสำรองปี 2011 (ล้านดอลลาร์)	ลำดับในดัชนีพัฒนาการตามบุคลากรของ UNDP
อินโดนีเซีย	242.3	846.7	6.5	3,511	203.6	24	110.1	124
ไทย	69.5	346.1	4	5,117	226.4	65.4	175.1	103
มาเลเซีย	28.9	288.1	5.1	9,949	227.1	78.8	133.6	61
สิงคโปร์	5.2	260	4.9	46,241	409.5	157.5	237.7	26
ฟิลิปปินส์	94.9	224.8	3.9	2,346	48.3	21.5	75.3	112
เวียดนาม	88.8	123.6	5.9	1,375	96.9	78.4	12.2	128
พม่า	48.3	51.9	10.4	857	0.1	0.1	-	149
บรูไน	0.4	16.4	2.2	38,715	12.4	76	1.7	33
ติมูทอ	14.3	12.9	6.9	909	5.3	41	3.4	139
ลาว	6.3	8.1	8	1,281	1.9	22.8	-	138
รวมเฉลี่ย	599	2,179	5.8	3,638	1,231	58.4	749	

<http://www.thai-aec.com/700> องค์ความรู้: ประชาคมเศรษฐกิจอาเซียน

23

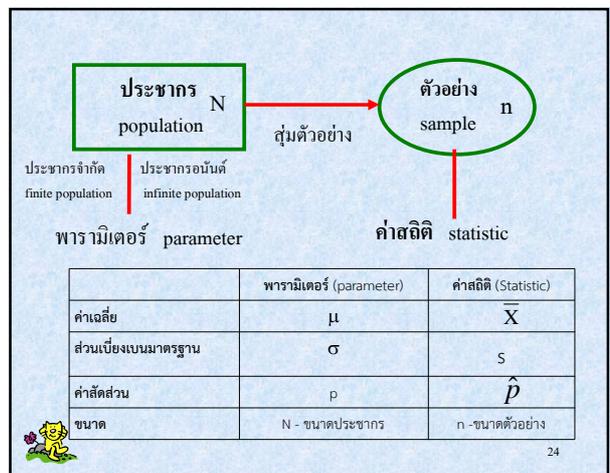
3. การรวบรวมจากแหล่งข้อมูลที่ได้เก็บรวบรวมข้อมูลไว้ (Published data)

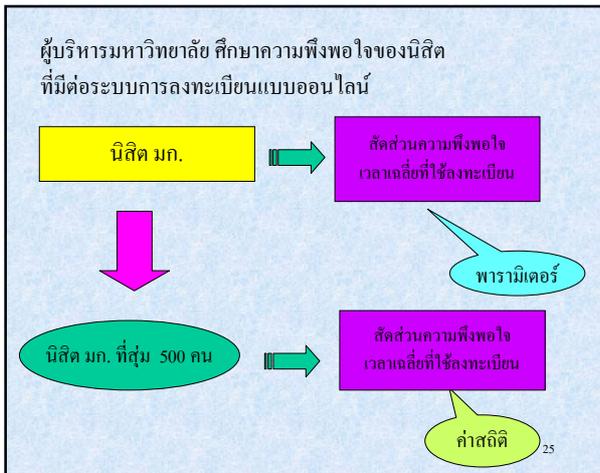
เอกสาร วารสารของหน่วยงานของรัฐ ธนาคาร สำนักงานสถิติ หนังสือทั่วไป ตำรา คู่มือ หนังสืออ้างอิง รายงานการวิจัย บทความวิจัย วิทยานิพนธ์วารสาร จุลสาร เอกสารทางวิชาการ ไมโครฟิล์ม เอกสารอื่นๆ

Website ของหน่วยงานต่างๆ

ข้อมูลทุติยภูมิ (secondary data)

21



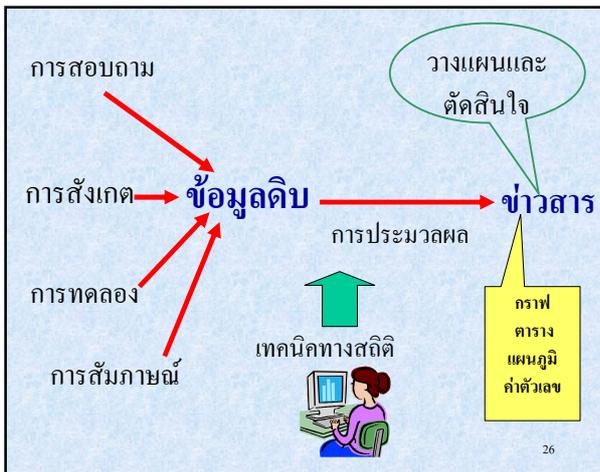


๑ การนำเสนอข้อมูล

นำข้อมูลที่เก็บรวบรวมมาจัดหมวดหมู่ เพื่อให้ผู้อ่านเข้าใจใน
รายละเอียดของข้อมูล และลักษณะสำคัญ เครื่องมือที่ใช้คือ

1. ตาราง ตารางทางเดียว ตารางสองทาง ตารางหลายทาง
2. กราฟ : กราฟแท่ง กราฟเส้น กราฟวงกลม ฮิสโตแกรม
แผนภาพกล่อง แผนภาพลำต้นและใบ

28



การนำเสนอข้อมูลด้วยตาราง

1. ตารางทางเดียว
ข้อมูลเชิงปริมาณ

ตัวอย่าง จำนวนสมาชิกของครอบครัว 40 ครอบครัว มีดังนี้

5	6	4	5	2	7	5	6	5	3
6	5	4	7	4	3	7	5	7	4
3	6	8	4	6	5	4	3	2	5
5	7	5	8	6	4	2	4	6	5

29

๓ การประมวลผลเชิงพรรณนา

เพื่ออธิบายลักษณะข้อมูล โดยการใช้กราฟ
ตาราง แผนภูมิ การคำนวณค่าเบื้องต้น โดย
ใช้สถิติเชิงพรรณนา (Descriptive Statistics)

- ๑ การนำเสนอข้อมูล
- ๒ ตำแหน่งข้อมูล
- ๓ ค่ากลาง
- ๔ ค่าวัดการกระจาย

27

การแจกแจงความถี่

1. การแจกแจงแบบไม่จัดกลุ่ม

จำนวนบุตร	รอยคะแนน	ความถี่
2		3
3		4
4		8
5		11
6		7
7		5
8		2
รวม		30

30

2. การแจกแจงแบบจัดกลุ่ม

ตัวอย่าง อายุ (ปี) ของลูกค้าห้างสรรพสินค้าแห่งหนึ่งจำนวน 50 คน มีดังนี้

21	22	58	12	55	42	32	66	51	72
45	27	57	44	65	37	59	44	37	56
49	45	36	45	28	62	41	57	71	63
35	29	47	40	39	65	16	43	52	50
24	48	53	77	36	55	51	32	48	35

จงสร้างตารางแจกแจงความถี่ ให้มีจำนวนชั้นเท่ากับ 7 ชั้น

31

2. ตารางสองทาง

ข้อมูลเชิงคุณภาพ

ตาราง ความคิดเห็นเกี่ยวกับการลงทะเบียนแบบเหมาจ่ายจำแนกตามเพศ

เพศ	ความคิดเห็น			รวม
	เห็นด้วย	ไม่เห็นด้วย	ไม่มีความเห็น	
ชาย	70	30	20	120
หญิง	100	60	20	180
รวม	170	90	40	300

34

จัดข้อมูลให้เป็นกลุ่มหรือชั้น (Class)

นับจำนวนข้อมูลในแต่ละ
กลุ่ม count - frequency

↓
ตารางแจกแจงความถี่
ข้อมูลเป็นช่วง

อายุ (ปี) (ขีดจำกัดชั้น)	รอยคะแนน	ความถี่
12 - 21		3
22 - 31		5
32 - 41		11
42 - 51		14
52 - 61		9
62 - 71		6
72 - 81		2
รวม		50

การแจกแจงความถี่ (Frequency distribution)

32

2. ตารางหลายทาง

ข้อมูลเชิงคุณภาพ

ตารางที่ 1 ความคิดเห็นของนักเรียนที่มีต่อการสอบเข้ามหาวิทยาลัยจำแนกตามสายและเพศ

ความคิดเห็น	วิทยาศาสตร์		สังคมศาสตร์		รวม
	ชาย	หญิง	ชาย	หญิง	
เห็นด้วย	80	90	60	100	330
ไม่เห็นด้วย	70	30	50	50	200
ไม่มีความเห็น	10	20	10	30	70
รวม	160	140	120	180	600

35

ตารางทางเดียวสำหรับข้อมูลเชิงคุณภาพ

ตารางที่ 1 จำนวนและร้อยละของสมาชิก
จำแนกตามระดับการศึกษา

ตารางที่ 1 จำนวนสินค้าที่ผลิตได้
จำแนกตามคุณภาพของสินค้า

คุณภาพสินค้า	จำนวน
เกรดดีมาก (A)	20
เกรดดี (B)	14
เกรดพอใช้ (C)	6

ระดับการศึกษา	จำนวนสมาชิก	ร้อยละ
ไม่ได้เรียน	10	25.0
ประถมศึกษา	15	37.5
มัธยมศึกษา	9	22.5
สูงกว่ามัธยมศึกษา	6	15.0
รวม	40	100.0

33

การนำเสนอข้อมูลด้วยกราฟ

ข้อมูลเชิงคุณภาพ

กราฟแท่ง

กราฟวงกลม

ข้อมูลเชิงปริมาณ

กราฟเส้น กราฟแท่ง

ฮิสโตแกรม

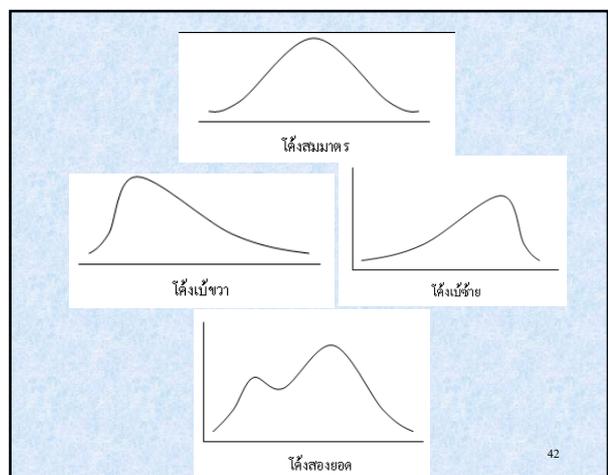
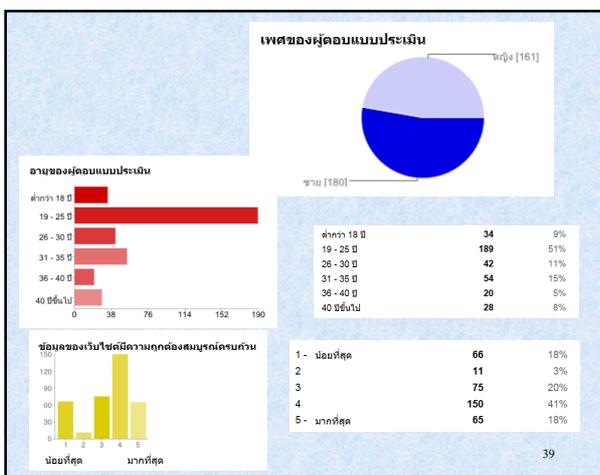
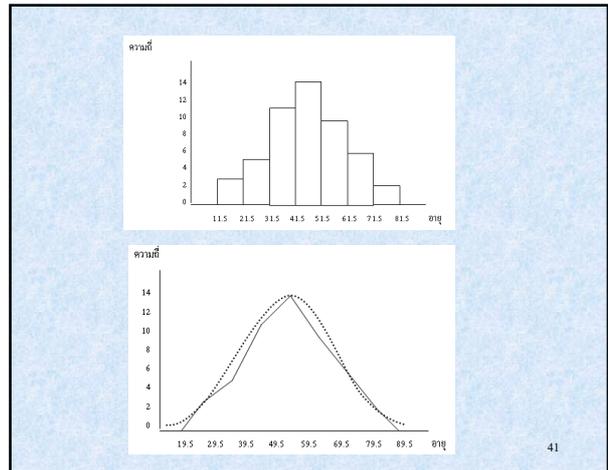
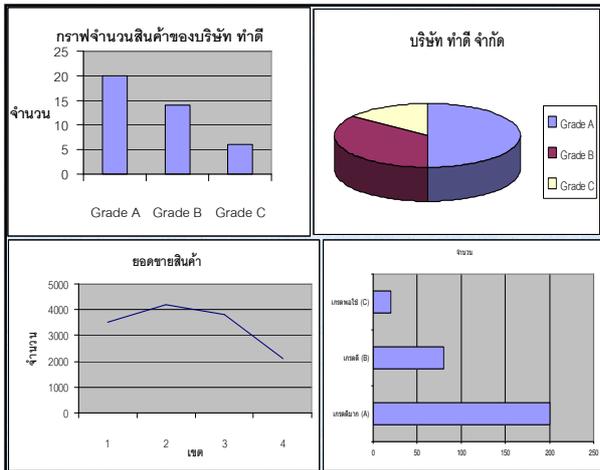
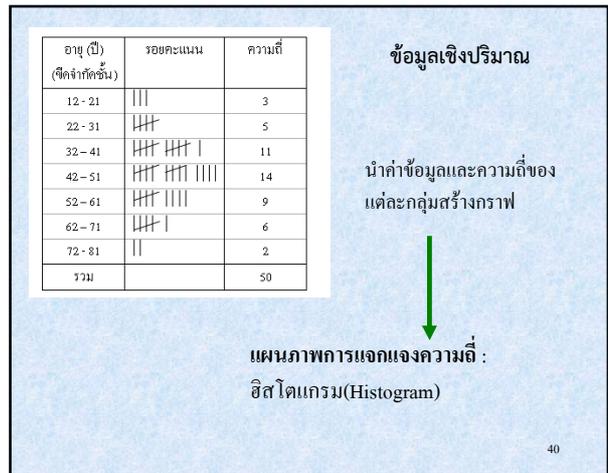
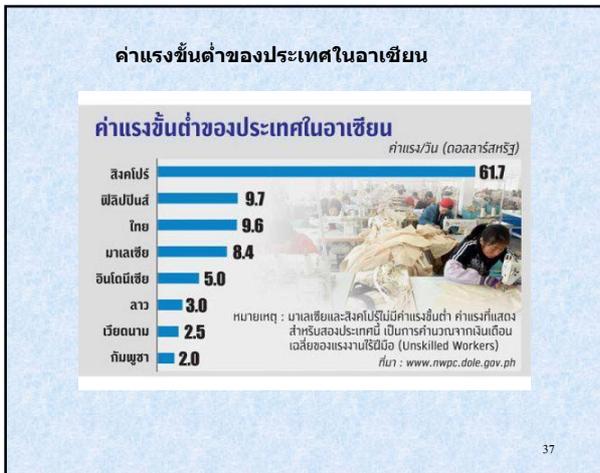
โค้งความถี่

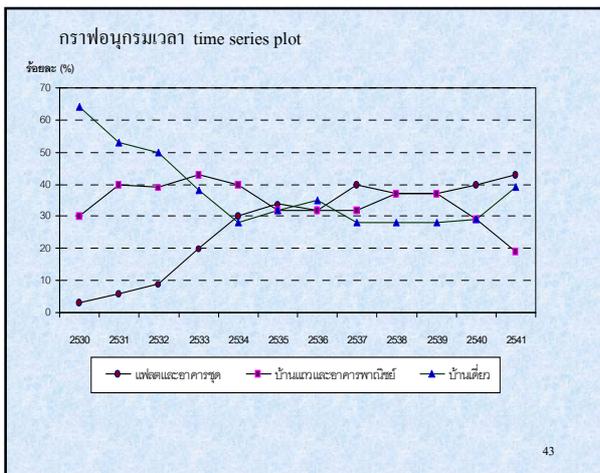
กราฟกล่อง

กราฟลำต้นและใบ

แผนภาพการกระจาย

36





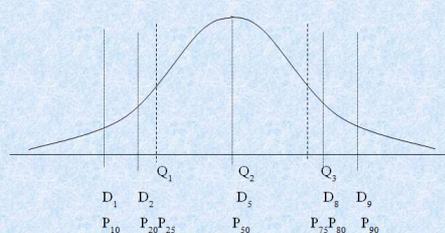
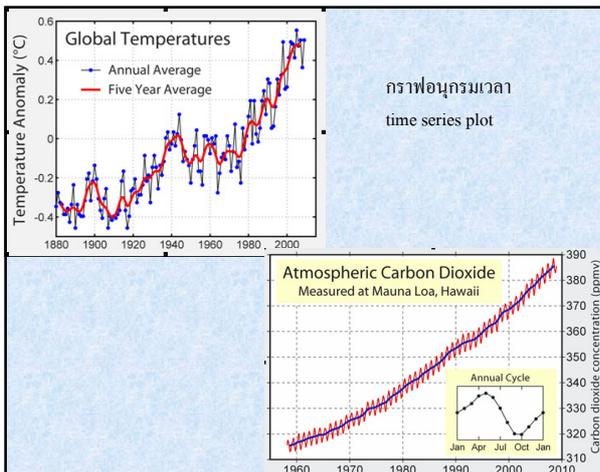
๑ ตำแหน่งข้อมูล

ควอไทล์ เป็นค่าข้อมูลที่อยู่ในตำแหน่งต่างๆ เมื่อแบ่งข้อมูลเป็น 4 ส่วนเท่าๆ กัน เรียกค่าที่อยู่ในแต่ละตำแหน่งนั้นว่า ควอไทล์ที่ 1 ควอไทล์ที่ 2 และ ควอไทล์ที่ 3 หรือ Q_1, Q_2 และ Q_3

เดซิล์ เป็นค่าข้อมูลที่อยู่ในตำแหน่งต่างๆ เมื่อแบ่งข้อมูลทั้งหมดเป็น 10 ส่วน (ตำแหน่ง) เรียกว่า เดซิล์ $D_1, D_2, D_3, \dots, D_9$

เปอร์เซ็นต์ไทล์ เป็นค่าข้อมูลที่อยู่ในตำแหน่งต่างๆ เมื่อแบ่งข้อมูลทั้งหมดเป็น 100 ส่วน (ตำแหน่ง) เรียกว่า เปอร์เซ็นต์ไทล์ $P_1, P_2, P_3, \dots, P_{99}$

46

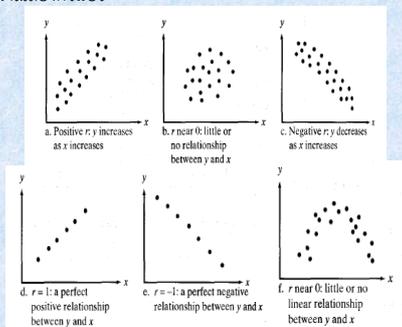


$$Q_1 = P_{25} \quad Q_2 = D_5 = P_{50} \quad Q_3 = P_{75}$$

$$D_1 = P_{10} \quad D_2 = P_{20} \dots D_9 = P_{90}$$

47

แผนภาพการกระจาย scatter plot แสดงความสัมพันธ์ระหว่างตัวแปร
เชิงปริมาณสองตัวแปร



อ้างอิงจาก A First course in business Statistics eighth edition: McClave, James T. Benson P.George and Sincich Terry. Prentice Hall.

45

การคำนวณค่าเปอร์เซ็นต์ไทล์ เดซิล์ และควอไทล์

$$Q_k = X_{\frac{k(n+1)}{4}} \quad k=1,2,3$$

$$D_k = X_{\frac{k(n+1)}{10}} \quad k=1,2,3,\dots,9$$

$$P_k = X_{\frac{k(n+1)}{100}} \quad k=1,2,3,\dots,99$$

48

แผนภาพกล่อง (Box and Whisker plot)

Lower extreme

Lower quartile

Median

Upper quartile

Upper extreme

Mean

ลักษณะของแผนภาพ : ประกอบด้วยค่าควอไทล์ (Q_1 , Q_2 และ Q_3) เป็นกล่องที่มีค่า Q_1 และ Q_3 เป็นค่าต่ำและค่าสูงของกล่องที่วางตามแนวนอนหรือแนวตั้ง

เป็นแผนภาพที่ใช้สรุปลักษณะของข้อมูล แผนภาพกล่องแสดงลักษณะที่สำคัญของข้อมูลชุดนั้น เช่น ค่ากลาง ค่าวัดการกระจาย ลักษณะการแจกแจงข้อมูล และค่าผิดปกติของข้อมูล (outlier)

49

ตาราง ข้อมูลระดับคลอเรสเตอรอลที่จัดเรียงลำดับแล้ว

ตารางที่ 5.2 ข้อมูลระดับคลอเรสเตอรอลที่จัดเรียงลำดับแล้ว

520	358	330	302	275	264	245	210
474	357	328	298	274	261	244	208
453	355	327	298	274	261	244	208
428	353	322	296	274	260	243	206
420	353	322	295	273	260	243	206
420	352	320	294	273	259	243	203
416	352	320	292	273	259	240	195
403	348	320	290	273	259	239	193
403	348	319	290	271	258	239	192
394	346	317	286	271	257	238	190
391	344	317	286	270	256	235	187
390	343	315	285	269	254	235	187
386	341	314	285	269	254	230	185
386	341	314	284	269	253	229	178
384	341	312	283	269	253	227	178
383	341	312	283	268	253	225	178
378	337	311	282	268	252	224	176
370	337	311	282	267	252	222	173
370	336	310	281	267	252	222	173
369	336	307	280	266	251	219	172
367	336	305	279	266	250	218	172
365	334	304	279	266	248	218	166
365	333	302	278	264	248	216	156
363	333	302	277	264	246	214	156
360	332	302	277	264	246	214	135

ข้อมูล

ความยาวของกล่อง $IQR = Q_3 - Q_1$
กลุ่มข้อมูล 50% ของทั้งหมด

ข้อมูลผิดปกติ (Outlier)

50

ตัวอย่าง แผนภาพกล่องของข้อมูลระดับคลอเรสเตอรอล

ระดับคลอเรสเตอรอล

$Q_1 = 245.5$ $Q_2 = 276$ $Q_3 = 331$

ความยาวกล่อง (IRQ) $= Q_3 - Q_1$
 $= 331 - 245.5$
 $= 85.5$

มีขบวนการ
มีข้อมูล
ผิดปกติค่าสูง

53

Negatively Skewed

Positively Skewed

Symmetric

51

กราฟกล่องแสดงข้อมูลในเชิงเปรียบเทียบความแตกต่าง

Boxplot of Height by Sex

Height

Sex

54

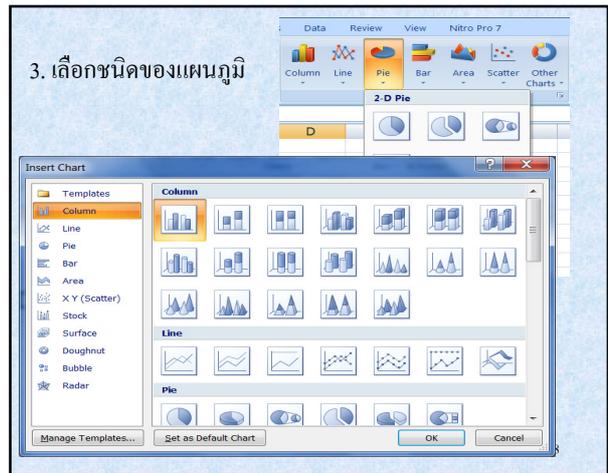
แผนภาพลำต้นและใบ Stem and leaf plot

ตัวอย่าง กราฟลำต้นและใบของคะแนนสอบของนักเรียนจำนวน 50 คน

75	98	42	75	84	87	65	59	63
86	78	37	99	66	90	79	80	89
68	57	95	55	79	88	76	60	77
49	92	83	71	78	53	81	77	58
93	85	70	62	80	74	69	90	62
84	64	73	48	72				

3	7
4	289
5	35789
6	022345689
7	01234556778899
8	00134456789
9	0023589

55

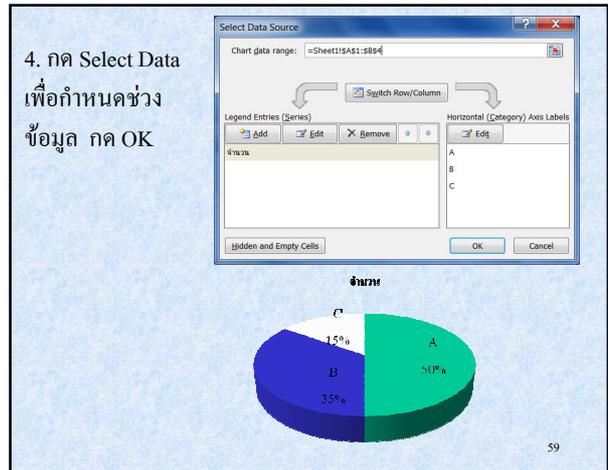


การนำเสนอข้อมูลด้วยโปรแกรม Excel

1. บันทึกข้อมูลลงใน Excel

	A	B	C
1	คุณภาพสินค้า	จำนวน	
2	A	20	
3	B	14	
4	C	6	

56



2. กดคำสั่งสร้างกราฟคือ insert

	A	B	C	D	E	F	G
1	คุณภาพสินค้า	จำนวน					Insert a pie chart.
2	A	20					Pie charts display the contribution of each value to a total.
3	B	14					Use it when values can be added together or when you have only one data series and all values are positive.
4	C	6					
5							
6							

57

ค่ากลาง

เพื่ออธิบายลักษณะข้อมูลด้วยค่าตัวเลข

ค่าวัดแนวโน้มสู่ส่วนกลางหรือค่ากลางที่สำคัญ คือ

- ค่าเฉลี่ย(mean)
- ค่ามัธยฐาน (median)
- ค่าฐานนิยม (mode)

60

ค่าเฉลี่ย (Average หรือ Mean)

$$\text{ค่าเฉลี่ย} = \frac{\text{ผลรวมของค่าของข้อมูลทั้งหมด}}{\text{จำนวนข้อมูลทั้งหมด}}$$

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

$$\mu = \frac{\sum_{i=1}^N X_i}{N}$$

ตัวอย่าง ข้อมูล 2 ชุดต่อไปนี้ คือ

ชุด A : 11 13 18 19 20

ชุด B : 4 4 9 11 12 14

ค่ามัธยฐานของข้อมูลชุด A คือ 18

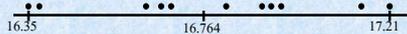
ค่ามัธยฐานของข้อมูลชุด B คือ $\frac{9 + 11}{2} = 10$

ตัวอย่าง ข้อมูลต่อไปนี้ เป็นระยะทางวิ่งต่อน้ำมัน 1 ลิตรของรถ
ขี้อี่ห้อหนึ่ง จำนวน 10 คัน ดังนี้ 16.85 16.4 17.21 16.35
16.52 17.04 16.96 17.15 16.59 และ 16.57

คำนวณระยะทางที่วิ่งได้เฉลี่ยต่อลิตร (\bar{X})

$$\bar{X} = \frac{16.85 + \dots + 16.57}{10} = 16.764$$

ค่าเฉลี่ยจะตกอยู่จุดสมมูลย์



ค่าฐานนิยม (Mode)

เป็นค่าของข้อมูลที่เกิดขึ้นบ่อยครั้งที่สุด

③ 6 9 ③ 5 8 ③ 10 4 6 ③ 1

Mode = 3

③ 6 9 ③ 5 8 ③ 10 4 6 ③ 1 6 2 5 1 6

Mode = 3 , 6 เรียกว่า Bimodal

ค่าฐานนิยมสามารถใช้ได้กับข้อมูลเชิงปริมาณและข้อมูล
เชิงคุณภาพ หรือไม่ ???

ค่ามัธยฐาน (median)

ค่ามัธยฐาน (median) เป็นค่ากลางที่แบ่งข้อมูลเป็น 2 ส่วนเท่าๆ
กันจะมีข้อมูลจำนวนครึ่งหนึ่งที่มีค่าต่ำกว่ามัธยฐาน และอีก
ครึ่งหนึ่งที่มีค่าสูงกว่ามัธยฐาน

$$\text{ค่ามัธยฐาน} = \frac{X_{\frac{n+1}{2}}}{2}$$

การหาค่ามัธยฐาน เรียงข้อมูลจำนวน n ค่าจากน้อยไปมาก

- ถ้า n เป็นจำนวนคี่ ค่ามัธยฐานคือ ค่าข้อมูลที่อยู่ตำแหน่งกึ่งกลาง
- ถ้า n เป็นจำนวนคู่ ค่ามัธยฐานคือ ค่าเฉลี่ยระหว่างข้อมูล 2 ค่าที่อยู่
ตำแหน่งกึ่งกลาง

ตัวอย่าง การใช้ฐานนิยมในข้อมูลเชิงคุณภาพ

เกรด B เป็นฐานนิยมของเกรดวิชานี้

สมชาย เป็นฐานนิยมของชื่อชายไทย

เลือดกลุ่ม O เป็นฐานนิยมของกลุ่มเลือดของคน

4 คำวัดการกระจายหรือความผันแปร (Measure of Variability)

ข้อมูล 3 ชุดคือ

ชุด 1 : 10 10 10 10 10

ชุด 2 : 9 11 11 9 10

ชุด 3 : 8 12 10 12 8

ค่าเฉลี่ยของข้อมูลทั้ง 3 ชุด = 10

แต่ข้อมูลทั้ง 3 ชุด มีค่าการกระจายที่ไม่เท่ากัน ดังนั้นควรใช้ทั้งค่า
กลางและคำวัดการกระจายในการอธิบายข้อมูล

67

ตัวอย่าง ข้อมูลต่อไปนี้เป็นระยะทางวิ่งต่อน้ำมัน 1 ลิตรของรถ
ซีห้อหนึ่ง จำนวน 10 คัน ดังนี้ 16.85 16.4 17.21 16.35
16.52 17.04 16.96 17.15 16.59 และ 16.57

จงคำนวณพิสัยและพิสัยระหว่างควอไทล์

$$\begin{aligned} \text{พิสัย} &= \text{ข้อมูลค่าสูงสุด} - \text{ค่าต่ำสุด} \\ &= 17.21 - 16.35 = 0.86 \end{aligned}$$

70

ค่าสถิติที่ใช้วัดการกระจาย หรือความผันแปร

- ❶ ค่าพิสัย (range)
- ❷ ค่าพิสัยระหว่างควอไทล์ (IQR)
- ❸ ค่าความแปรปรวน (Variance)
- ❹ ค่าเบี่ยงเบนมาตรฐาน (Standard deviation)

68

16.35 16.4 16.52 16.57 16.59 16.85 16.96 17.04 17.15 17.21

$$\begin{aligned} Q_1 &= X_{\frac{1(n+1)}{4}} = X_{\frac{1(10+1)}{4}} = X_{2.75} \\ &= 16.4 + 0.75(16.52 - 16.4) = 16.49 \end{aligned}$$

$$\begin{aligned} Q_3 &= X_{\frac{3(n+1)}{4}} = X_{\frac{3(10+1)}{4}} = X_{8.25} \\ &= 17.04 + 0.25(17.15 - 17.04) = 17.068 \end{aligned}$$

$$\begin{aligned} \text{พิสัยระหว่างควอไทล์ IQR} &= Q_3 - Q_1 \\ &= 17.068 - 16.49 = .577 \end{aligned}$$

71

- ❶ พิสัย range

$$\text{พิสัย} = \text{ข้อมูลค่าสูงสุด} - \text{ค่าต่ำสุด}$$

ค่าพิสัยมาก แสดงว่าข้อมูลกระจายหรือผันแปรมาก

- ❷ ค่าพิสัยระหว่างควอไทล์ Interquartile range

$$\text{IQR} = Q_3 - Q_1$$

IQR วัดการกระจายได้ดีกว่าค่าพิสัย เพราะเมื่อมีค่าสูงหรือต่ำ
เกินไป จะมีผลต่อ IQR น้อยกว่าพิสัย

69

- ❸ ค่าความแปรปรวนและ ❹ ค่าเบี่ยงเบนมาตรฐาน

เป็นค่าวัดความผันแปรหรือวัดการกระจายของข้อมูลที่นิยม
ใช้กันมาก เนื่องจากเป็นค่าที่ใช้ข้อมูลทุกค่ามาคำนวณ

$$\text{ค่าเบี่ยงเบนมาตรฐาน} = \sqrt{\text{ความแปรปรวน}}$$

72

X_1, X_2, \dots, X_N X_1, X_2, \dots, X_n

ความแปรปรวน (Variance)

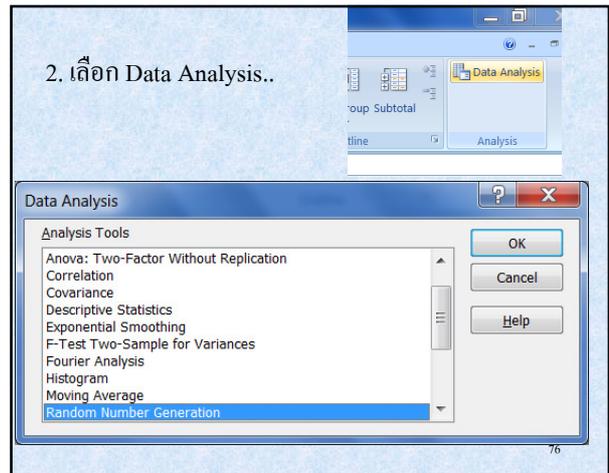
$$\sigma^2 = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N}$$

$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$$

$$\sigma^2 = \frac{\sum_{i=1}^N X_i^2}{N} - \mu^2$$

$$S^2 = \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{n - 1}$$

73



ตัวอย่าง ข้อมูลต่อไปนี้ เป็นระยะทางวิ่งต่อน้ำมัน 1 ลิตรของรถยี่ห้อหนึ่ง จำนวน 10 คัน ดังนี้ 16.85 16.4 17.21 16.35 16.52 17.04 16.96 17.15 16.59 และ 16.57 จงคำนวณส่วนเบี่ยงเบนมาตรฐาน

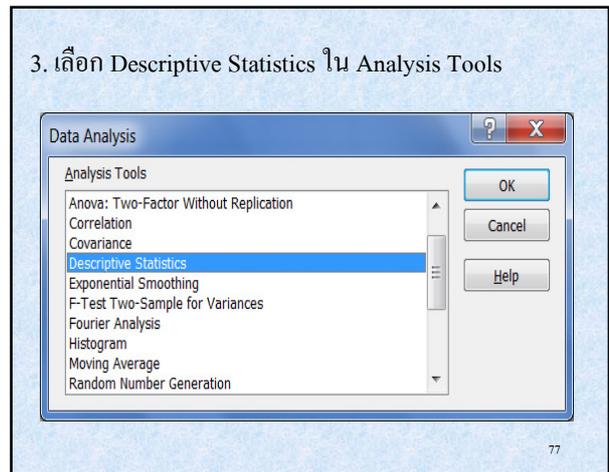
$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$$

$$= \frac{(16.85 - 16.764)^2 + (16.4 - 16.764)^2 + \dots + (16.57 - 16.764)^2}{10 - 1}$$

$$= .10013$$

ส่วนเบี่ยงเบนมาตรฐาน = $\sqrt{.10013} = .3164$

74



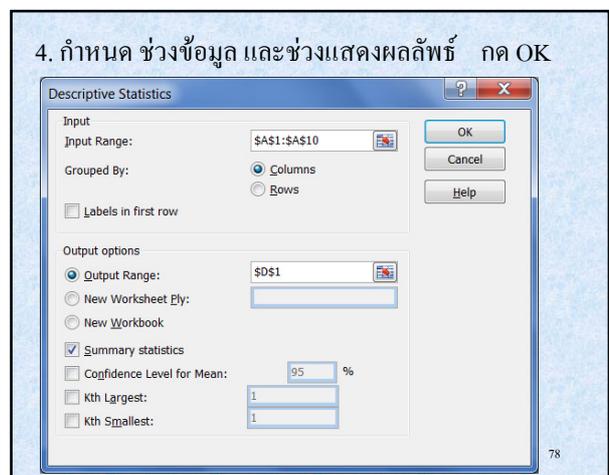
ตัวอย่าง การวิเคราะห์ข้อมูลสถิติโดยใช้คอมพิวเตอร์

ระยะทางที่รถวิ่งได้ต่อน้ำมัน 1 ลิตร จากรถ 10 คัน
16.85 16.4 17.21 16.35 16.52
17.04 16.96 17.15 16.59 16.57

1. บันทึกข้อมูลลงใน Excel ในคอลัมน์ A

	A	B
1	16.85	
2	16.4	
3	17.21	
4	16.35	
5	16.52	
6	17.04	
7	16.96	
8	17.15	
9	16.59	
10	16.57	
11		

75



ผลลัพธ์

C	D	E
	Column1	
Mean		16.764
Standard Error	0.100068865	
Median		16.72
Mode		#N/A
Standard Deviation	0.316445537	
Sample Variance	0.100137778	
Kurtosis	-1.68586664	
Skewness	0.12770493	
Range		0.86
Minimum		16.35
Maximum		17.21
Sum		167.64
Count		10

79

