




สวัสดีค่ะ

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

1

การอนุมานสถิติในชีวิตประจำวัน

- ❶ ความสำคัญของสถิติ
- ❷ ข้อมูลสถิติและการเก็บรวบรวมข้อมูล
- ❸ การประมวลผลข้อมูล
 - ❸.1 การประมวลผลเชิงพรรณนา
 - ❸.2 การประมวลผลเชิงอนุมาน




2

❶ ความสำคัญของสถิติ

ความหมายของสถิติ

1. ตัวเลข กลุ่มตัวเลขหรือตัวเลขสถิติ
2. ศาสตร์หรือวิชา การรวบรวมข้อมูล การจัดระเบียบข้อมูล การนำเสนอข้อมูล การวิเคราะห์ การแปลความหมายและตีความ การวิเคราะห์เชิงสถิติระดับสูง




3

❶ ความสำคัญของสถิติ

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

'Statistics is around'

สถิติในชีวิตประจำวันมีมากมายไม่จำกัดขอบเขตว่าอยู่ในเรื่องใดเรื่องหนึ่ง และการใช้สถิติอย่างไม่รู้ตัวอยู่เป็นประจำ



4

ราคาน้ำมันวันนี้

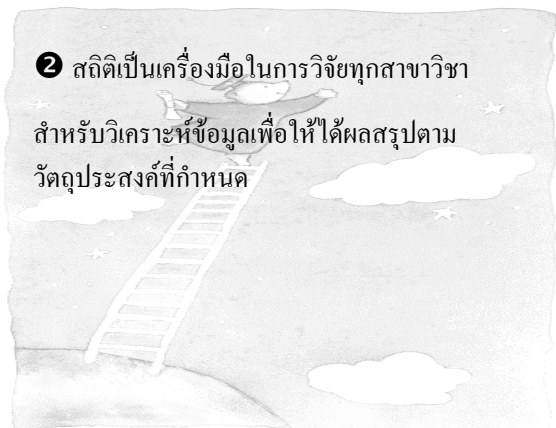
เบนซิน 91	36.84	ต่อลิตร
แก๊สโซฮอล์ 95	33.04	ต่อลิตร
แก๊สโซฮอล์ 91	31.54	ต่อลิตร
เบนซิน E20	30.74	ต่อลิตร
ดีเซลหมุนเร็ว	29.89	ต่อลิตร
ดีเซล บี5	28.69	ต่อลิตร
เอ็นจีวี	8.50	ต่อกก.

ราคากองค่าวันนี้

ราคาซื้อทองคำแท่ง	ซื้อบาทละ	18,300
ราคาขายทองคำแท่ง	ขายบาทละ	18,400
ราคาซื้อทองคำรูปพรรณ	ซื้อบาทละ	18,040.40
ราคาขายทองคำรูปพรรณ	ขายบาทละ	18,800

5

❷ สถิติเป็นเครื่องมือในการวิจัยทุกสาขาวิชา
สำหรับวิเคราะห์ข้อมูลเพื่อให้ได้ผลสรุปตาม
วัตถุประสงค์ที่กำหนด



ตัวอย่างงานวิจัยที่ใช้สถิติเป็นหลักฐานซึ่งเป็นข้อมูลเชิงประจักษ์และใช้วิธีการทางสถิติในการสรุปผลการวิจัย

ศึกษาวิจัยพบ ความสำคัญของการนอนนอนไม่ถึง 6 ชม. อายุสั้น
การนอนน้อยอาจจะเป็นสาเหตุของสุขภาพเสื่อมโดยตรง ซึ่งนำไปสู่การมีอายุสั้นในที่สุด นักวิจัยทางการแพทย์ทั้งเมือง มะกะโรนีและนักชา กล่าวเตือนคนที่อดหลับอดนอนว่า ผู้ที่นอนหลับไม่ ถึงคืนละ 6 ชม. มีหวังจะได้เค็ดสเมอเร็ก่อนวัยอันควร คนนอนน้อยเช่นนั้นอยู่เป็นประจำมีหวังอายุสั้น จะเสียชีวิตลงภายในช่วงเวลา 25 ปีข้างหน้ามากกว่าผู้ที่นอนนานตามเวลาที่เหมาะสม ระหว่าง 6-8 ชม. กว่ากันถึงร้อยละ 12 แต่ก็น่าแปลกใจเช่นกันที่พวกเขาจึงพบว่าผู้ที่ นอนนานเกิน 9 ชม. ก็เสี่ยงกับอายุสั้นด้วยเช่นกัน

7

วารสาร วิชาการ "การนอน" นักวิจัยของสองชาติได้ ร่วมกันศึกษา ความเกี่ยวพันของการนอนกับอัตราการเสียชีวิต ด้วยการทบทวน รายงานการศึกษาเกี่ยวกับการนอน 16 เรื่อง ซึ่งเกี่ยวพันกับผู้คนมากถึง 1,500,000 คน ได้ข้อสรุปว่า "การเสียชีวิตก่อนวัยอันควรทั้งหมด ล้วนแต่เกี่ยวพันกับการนอนน้อยไปหรือนอนมากไป มากน้อยเกินกว่าเวลานอนที่เหมาะสมที่ควร คืนละระหว่าง 6-8 ชม. ทั้งสิ้น" คณะนักวิจัยเชื่อว่า ขณะที่การนอนน้อยอาจจะเป็นสาเหตุของสุขภาพเสื่อมโดยตรง ซึ่งนำไปสู่การมีอายุสั้นในที่สุด แต่การนอนนานก็อาจเป็นเพียงเครื่องบอกให้รู้ว่า สุขภาพไม่ดี.

ข้อมูลข่าวโดยหนังสือพิมพ์ไทยรัฐออนไลน์ วันที่ 8 พฤษภาคม 2553

8

นักเรียนชอบกินฟาสต์ฟู้ดสมองจะเสื่อม ถึงกับสอบตกได้

ไทยรัฐออนไลน์
โดย หนึ่งข่าวรั่วอากาศ-ไอที
2 มิถุนายน 2552, 05:00 น.

tags:
นักเรียน, ชอบกิน, ฟาสต์ฟู้ด, สอบตก



เด็กนักเรียนระดับประถมศึกษา ที่กินอาหารฟาสต์ฟู้ด เช่น แฮมเบอร์เกอร์และมันฝรั่งทอดมากเกินไป อาจทำให้การเรียนเสื่อม ถึงกับสอบตกได้

หนังสือพิมพ์รายวัน "เดอะ เทเลกราฟ" เมืองซิดนีย์ออก เสนอข่าวว่า ได้มีการศึกษาเกี่ยวกับนักเรียนชั้นประถมศึกษาอเมริกัน จำนวน 5,500 คนพบว่า เด็กคนที่กินอาหารฟาสต์ฟู้ด อาหารจานเดียวมากกว่าสามมื้อ อย่างต่อเนื่องเป็นประจำจะสอบได้ต่ำ บางรายจะคะแนนวิชาอ่านออกเขียนได้ และวิชาเลขต่ำกว่าค่าเฉลี่ยโดยเฉลี่ยถึงร้อยละ 16

คณะนักวิจัยมหาวิทยาลัยแวนคูเวอร์ บริติชโคลัมเบีย ได้ศึกษาเกี่ยวกับการกินของเด็กนักเรียนอเมริกันชั้น 10-11 ขวบ เปรียบเทียบกับผลของการทดสอบวิชาอ่านออกเขียนได้และวิชาเลข หลังจากที่ได้ ทราบว่าเด็กเหล่านั้นกินครั้ง กินอาหารฟาสต์ฟู้ดมา เมื่ออาทิตย์ก่อน คืนละ 3 มื้อ

รายงานผลการศึกษานี้พบว่า โยบปกติแล้วเด็กจะทำการคะแนนในการทดสอบ ได้มากกว่าระหว่าง 58-181 คะแนน หรือคะแนนเฉลี่ย 14.5 คะแนน แต่เด็กที่กินอาหารฟาสต์ฟู้ดทุก อาทิตย์หนึ่ง ถึง 4-5 ครั้ง จะทำการคะแนนได้ต่ำ กว่าคะแนนเฉลี่ยเกือบ 7 คะแนน ยิ่งผู้ที่กินวันละ 1 ครั้ง จะทำการคะแนนได้ต่ำลง 16 คะแนน และยิ่งกินแต่กินวันละ 3 มื้อ จะทำการคะแนนต่ำลงถึง 19 คะแนน.

กินอาหารไขมันสูง อาจทำร่างกายหลังลดแถมความจำแย่ลง

โดย ASTVผู้จัดการออนไลน์ 9 กันยายน 2552 13:44 น.

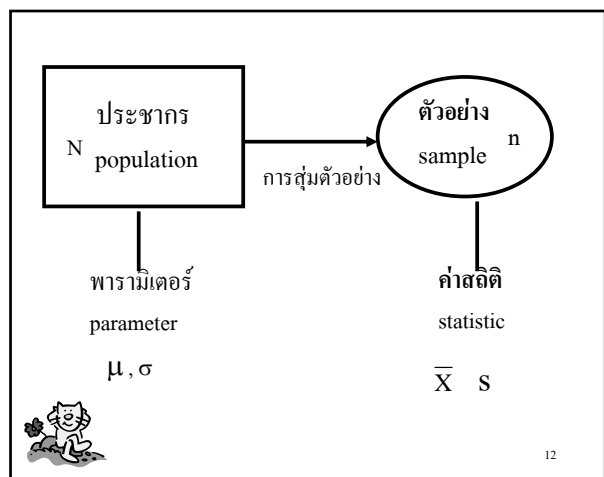
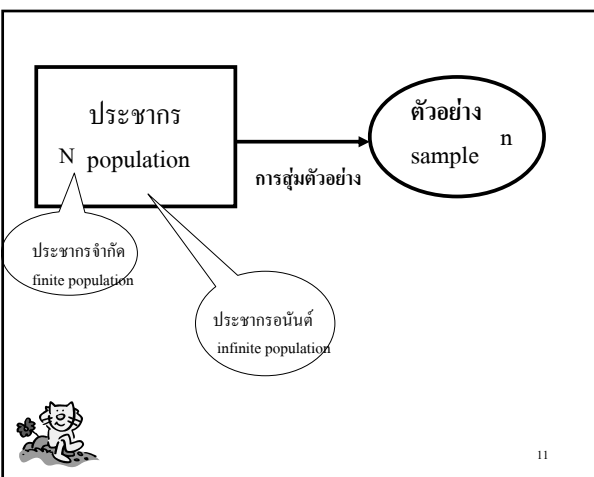
คลิกที่ภาพเพื่อดูขนาดใหญ่ขึ้น

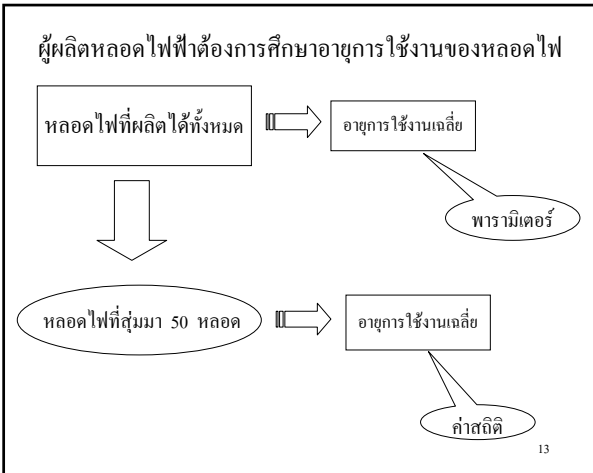


กินอาหารไขมันสูงอาจเสี่ยง เป็นโรคอ้วนหรือโรคหัวใจในระยะยาวได้ แต่ในระยะสั้นอาจทำให้คนเราก่อนกำลังกายได้น้อยลง แคมเปญความเสี่ยงเรื้อรัง หลังนักวิจัยอังกฤษทดลองเสี่ยงหยุดอาหารไขมันสูง พบหญิงได้ระยะทางสั้นลง และหลงทางใน ชาวอกมากกว่าขึ้น

ดร.แอนดรูว์ เมอร์เรย์ (Dr Andrew Murray) และทีมวิจัยที่มหาวิทยาลัยออกฟอร์ด (Oxford University) สหราชอาณาจักร ศึกษาผลของคาร์โบไฮเดรตที่มีไขมันสูงมากเกินไปในหนูทดลอง พบว่าทำให้หนูมีความอดทนทางร่างกายน้อยลง และความสามารถของระบบความจำแย่ลง ซึ่งไขมันดีระบุว่าผลงานวิจัยนี้ได้ตีพิมพ์ในวารสารเอฟเอเอสเอ็มอี (FASEB)

"เราพบว่าเมื่อเสี่ยงหยุดอาหารที่มีไขมันสูง แทนอาหารสูตรมาตรฐานแบบเดิมที่มีไขมันต่ำ ทำให้พวกมันมีการแสดงออกทางกายภาพลดลงอย่างรวดเร็วจนน่าตกใจมาก ซึ่งหลังจากผ่านไปแค่ 9 วัน พวกมันวิ่งได้ไกลเพียง 50% ของระยะทางที่พวกมันเคยวิ่งเมื่อครั้งที่ยังถูกเลี้ยงด้วยอาหารไขมันต่ำ" ดร.เมอร์เรย์ เผยผลการวิจัยดังกล่าว ซึ่ง





2 ข้อมูลและการเก็บรวบรวมข้อมูล

ข้อมูล (data) คือข้อเท็จจริงหรือรายละเอียดเกี่ยวกับค่าที่ต้องการสนใจศึกษา ตัวเลข ตัวอักษร หรือ สัญลักษณ์ ได้จากการนับ การวัด การสังเกตหรือการสัมภาษณ์

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

ประเภทของข้อมูล

1. ข้อมูลเชิงปริมาณหรือข้อมูลจากการวัด

Quantitative data, Measurement data, Numerical data เป็นข้อมูลที่สามารถวัดได้เป็นตัวเลขและมีหน่วยแน่นอน

น้ำหนัก ความสูง อายุการใช้งานของหลอดไฟ ผลผลิตของข้าว ค่าธรรมเนียมต่างๆ ค่าใช้จ่าย

ข้อมูลต่อเนื่อง (continuous data) เป็นข้อมูลตัวเลขที่มีค่าติดต่อกันเป็นช่วง เช่น น้ำหนัก คะแนนสอบ

ข้อมูลไม่ต่อเนื่อง (discrete data) ได้จากการนับตามสิ่งที่น่าสนใจ เช่น จำนวนคนในครอบครัว จำนวนเสียงโทรศัพท์ที่เข้ามาในสำนักงาน จำนวนบุตรในครอบครัว จำนวนครั้งที่นิสิตขาดเรียน

2. ข้อมูลเชิงคุณภาพ qualitative data

เป็นข้อมูลที่ไม่สามารถวัดเป็นตัวเลขได้ แต่แสดงข้อเท็จจริงตามคุณสมบัติ

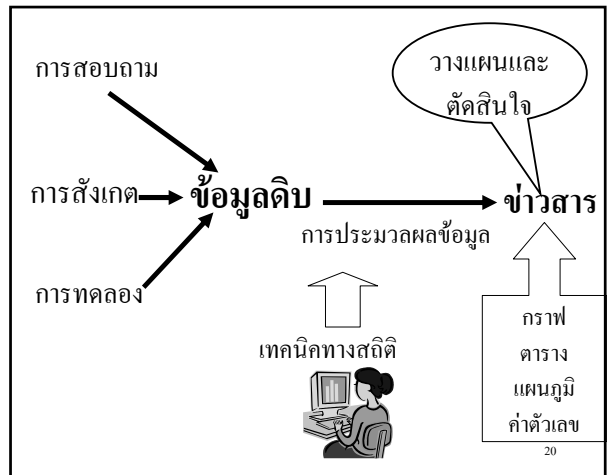
คณะ – เกษตร ประมง วิทยาศาสตร์ ...
 เพศ - ชาย หญิง
 วุฒิการศึกษา – ป.ตรี ป.โท ป.เอก
 ชั้นปี – 1 2 3 4
 ความพอใจต่อระบบการลงทะเบียนแบบออนไลน์ของมหาวิทยาลัย – พอใจ ไม่พอใจ

3. ข้อมูลจำแนกประเภท categorical data

เป็นข้อมูลที่นำมาจำแนกตามประเภทชนิดหรือลักษณะ โดยนับจำนวนในแต่ละกลุ่มเป็นความถี่

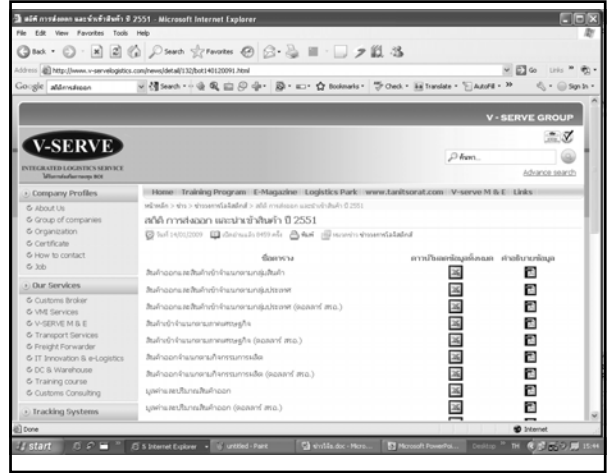
ข้อมูลจำนวนนิสิตในมก. จำแนกตามคณะ
 ข้อมูลจำแนกคนตามกลุ่มเลือด
 ข้อมูลจำนวนรถยนต์จำแนกตามยี่ห้อ
 ข้อมูลจำนวนนิสิตของมก. จำแนกตามคณะ
 ข้อมูลจำนวนนิสิตของมก. จำแนกตามเพศ

คณะ/โรงเรียนวิชา	เอก	โท	ดุษฎี	รวม
1 คณะเกษตร	15	121	197	343
2 คณะบริหารธุรกิจ	0	502	493	995
3 คณะวิศวกรรม	7	40	129	176
4 คณะอุตสาหกรรม	0	57	351	408
5 คณะวนศาสตร์	1	53	166	219
6 คณะวิทยาศาสตร์	9	129	323	461
7 คณะสัตวแพทยศาสตร์	0	521	947	1468
8 คณะศึกษาศาสตร์	31	376	199	606
9 คณะศิลปกรรมศาสตร์	2	323	438	763
10 คณะศึกษาศาสตร์	0	323	469	793
11 คณะสิ่งแวดล้อมและทรัพยากรศาสตร์	0	14	88	102
12 คณะเทคโนโลยีการเกษตร	1	65	341	407
13 คณะสัตวแพทยศาสตร์	1	27	57	85
14 คณะอุตสาหกรรมเกษตร	9	115	252	376
15 คณะวิทยาศาสตร์และเทคโนโลยี	1	55	241	337
16 คณะวิศวกรรมศาสตร์ ศึกษาศาสตร์	1	10	289	300
17 คณะสัตวแพทยศาสตร์และสัตวศาสตร์	0	57	257	314
18 คณะวิทยาศาสตร์และเทคโนโลยี	0	0	43	43
19 คณะวิทยาศาสตร์และเทคโนโลยี	0	16	0	16
20 คณะวิทยาศาสตร์และเทคโนโลยี	0	0	850	850
21 คณะวิทยาศาสตร์และเทคโนโลยี	0	0	132	132
22 คณะวิทยาศาสตร์และเทคโนโลยี	0	0	231	231
23 วิทยาลัยนานาชาติ สาขาวิชา ศึกษาศาสตร์	0	0	57	57
24 วิทยาลัยนานาชาติ สาขาวิชา ศึกษาศาสตร์	0	128	0	128
25 คณะวิทยาศาสตร์และเทคโนโลยี	0	0	104	104
26 คณะวิทยาศาสตร์และเทคโนโลยี	0	0	229	229
27 คณะวิทยาศาสตร์และเทคโนโลยี	0	368	254	622
28 วิทยาลัยนานาชาติ สาขาวิชา ศึกษาศาสตร์	0	0	81	81
29 วิทยาลัยนานาชาติ สาขาวิชา ศึกษาศาสตร์	0	0	54	54
30 วิทยาลัยนานาชาติ สาขาวิชา ศึกษาศาสตร์	5	15	0	20
31 วิทยาลัยนานาชาติ สาขาวิชา ศึกษาศาสตร์	0	0	344	344
32 วิทยาลัยนานาชาติ สาขาวิชา ศึกษาศาสตร์	0	0	54	54
33 วิทยาลัยนานาชาติ สาขาวิชา ศึกษาศาสตร์	0	3	61	64
34 วิทยาลัยนานาชาติ สาขาวิชา ศึกษาศาสตร์	0	0	107	107
รวม	186	2041	7197	10324



การเก็บรวบรวมข้อมูล


1. การสำรวจ
 - การสำมะโน (census) สำรวจทั้งหมด
 - การสุ่มตัวอย่าง (sampling survey) สำรวจบางส่วน
 - 'ข้อมูลจากการสำรวจ (Survey data)'
2. การทดลอง กำหนดสถานการณ์หรือการทดลอง
 - 'ข้อมูลที่ได้ออกจากการทดลอง (Experimental data)'
3. การรวบรวมจากแหล่งข้อมูลที่ได้เก็บรวบรวมข้อมูลไว้
 - หน่วยงานของรัฐ ธนาคาร สำนักกานสถิติ



	FD	FE	FF	FG	FH	FI	FJ	FK	FL	FM
จำนวนประชากร	384,266.53	426,370.17	381,987.95	439,374.85	483,770.82	497,000.56	468,876.56	485,963.23	455,632.56	385,099.24
จำนวนชาย	19,920.01	15,028.00	14,022.00	15,442.00	15,791.03	17,490.16	16,819.70	16,864.52	15,714.54	13,059.00
จำนวนหญิง	8,255.01	7,407.46	6,529.88	8,495.26	8,939.21	10,621.95	9,695.84	8,613.30	6,792.32	7,982.79
จำนวนเด็ก	3,446.80	4,116.12	3,737.89	3,897.68	3,897.84	4,959.89	3,895.69	3,827.11	3,442.85	2,453.79
จำนวนผู้ใหญ่	3,212.49	3,902.42	3,295.90	3,179.74	3,295.47	4,498.20	3,756.24	3,614.10	3,568.52	2,564.88
จำนวนเด็กชาย	4,328.20	4,305.28	3,328.13	4,580.01	4,547.25	4,876.25	5,129.41	4,778.75	4,543.29	4,188.88
จำนวนเด็กหญิง	954.92	966.20	826.89	960.09	956.52	1,080.95	1,119.05	1,077.65	916.13	816.13
จำนวนชายโต	171.81	148.77	157.80	169.41	171.52	206.47	215.22	212.22	239.42	172.03
จำนวนชายโต	927.79	962.01	826.53	907.70	926.56	1,024.54	934.68	921.46	878.18	816.13
จำนวนชายโต	47.01	47.96	59.84	68.24	76.47	84.25	82.59	82.72	72.15	48.74
จำนวนเด็กชายโต	2,289.65	2,229.54	2,298.86	2,403.61	2,466.16	2,618.76	2,897.96	2,617.00	2,442.99	2,170.88
จำนวนเด็กชายโต	65,690.43	73,094.96	62,627.07	63,736.75	74,899.58	73,644.37	75,000.82	77,895.07	74,791.13	61,467.61
จำนวนเด็กชายโต	26,374.89	37,263.96	34,605.85	36,301.80	47,536.16	39,512.20	40,664.47	43,886.24	41,236.37	32,493.77
จำนวนเด็กชายโต	3,079.23	16,522.11	6,138.14	8,988.40	6,846.76	8,528.24	9,464.20	9,399.60	8,524.10	7,893.11
จำนวนเด็กชายโต	21,245.30	29,388.98	20,993.07	23,916.75	26,975.05	25,493.30	24,531.95	24,843.12	27,668.64	21,433.79
จำนวนเด็กชายโต	14,077.04	27,842.99	11,776.96	19,251.14	25,451.98	20,979.72	14,886.25	31,771.82	35,265.46	17,447.01
จำนวนเด็กชายโต	37,521.74	46,101.19	37,965.70	43,482.79	48,774.76	45,504.29	47,479.76	49,979.89	44,991.85	33,521.38
จำนวนเด็กชายโต	15,968.25	16,373.00	16,303.64	19,280.90	21,625.26	19,834.17	20,527.56	21,564.32	19,847.71	17,796.43
จำนวนเด็กชายโต	2,181.57	2,098.01	1,489.45	1,715.56	1,682.90	1,981.27	1,919.24	1,994.42	1,722.67	1,240.38
จำนวนเด็กชายโต	13,227.79	4,261.61	3,254.44	3,899.09	4,698.72	4,645.14	4,288.56	4,493.52	6,376.61	3,796.63
จำนวนเด็กชายโต	16,638.01	19,366.25	16,581.25	18,796.83	20,993.77	19,763.79	20,743.60	21,045.04	19,028.85	15,229.89
จำนวนเด็กชายโต	2,280.35	3,997.17	2,997.62	2,979.15	3,989.80	3,660.62	3,589.61	3,232.90	2,957.33	2,874.88
จำนวนเด็กชายโต	33,725.01	37,022.53	31,538.87	35,933.73	42,621.40	35,366.58	34,018.84	36,366.77	34,000.53	27,337.15
จำนวนเด็กชายโต	11,687.01	12,544.98	10,675.01	10,823.80	9,748.41	9,532.03	9,593.15	7,598.87	7,294.26	6,895.75
จำนวนเด็กชายโต	3,534.88	3,846.02	3,208.88	3,905.98	4,056.80	4,951.09	3,982.88	4,347.35	4,114.17	3,248.43
จำนวนเด็กชายโต	1,414.79	1,488.68	1,142.57	1,164.46	1,088.54	1,209.51	1,203.54	1,570.11	1,326.35	1,021.72
จำนวนเด็กชายโต	3,279.08	2,357.87	3,141.23	4,467.47	4,959.87	5,539.89	5,931.15	6,486.92	4,291.53	3,031.11
จำนวนเด็กชายโต	2,626.85	2,727.29	2,172.88	2,432.72	2,181.86	2,299.54	2,216.47	2,451.63	2,898.99	2,295.20
จำนวนเด็กชายโต	148.04	142.15	191.60	238.15	194.64	139.78	189.98	143.94	145.54	95.00
จำนวนเด็กชายโต	11,622.22	12,005.90	10,807.10	12,073.89	11,211.12	12,796.72	13,084.73	16,657.43	14,894.19	11,836.89

กลุ่มตัวอย่างจากลูกค้า 400 คน แบ่งเป็น 4 กลุ่มๆละ 100 คน กำหนดให้ใช้แซมพุระผสม 4 ชั้นหรือ สอบถามความพอใจในการใช้แซมพุทั้ง 4 ชั้น

เป็นการศึกษาจากการทดลอง เพราะมีการควบคุมปัจจัยจำนวน 1 ปัจจัยคือ แซมพุที่ถูกใช้กำหนด 4 ชั้น




กลุ่มสอบถามประชาชน 1000 คน เกี่ยวกับจำนวนครั้งในการชมภาพยนตร์ในโรงภาพยนตร์ต่อเดือน รายได้ อายุ เพศ การศึกษา

เป็นการศึกษาจากการสังเกตและสอบถามจากสภาพจริง เพราะไม่มีการควบคุมปัจจัยที่อาจส่งผลกระทบต่อข้อมูลที่รวบรวม

25

3.1 การประมวลผลเชิงพรรณนา

เพื่ออธิบายลักษณะข้อมูล โดยใช้กราฟ ตาราง แผนภูมิ การคำนวณค่าเบื้องต้น โดยใช้สถิติเชิงพรรณนา (Descriptive Statistics)



- 1 การนำเสนอข้อมูล
- 2 ตำแหน่งข้อมูล
- 3 ค่ากลาง
- 4 ค่าวัดการกระจาย

26

1 การนำเสนอข้อมูล

นำข้อมูลที่เก็บรวบรวมมาจัดหมวดหมู่ เพื่อให้ผู้อ่านเข้าใจในรายละเอียดของข้อมูล และลักษณะสำคัญ เครื่องมือที่ใช้คือ

1. แผนภาพ : แผนภูมิแท่ง ฮิสโตแกรม แผนภาพกล่อง แผนภาพลำต้นและใบ กราฟวงกลม กราฟแท่ง กราฟเส้น ฯลฯ
2. ตารางแจกแจงความถี่

27

การฝึกสมาธิกับความจำ

Abrams, University of California

28

รายได้ของคนแบ่งเป็นกลุ่มรายได้สูง ปานกลาง ต่ำ

ข้อมูลเชิงปริมาณ

ข้อมูลเชิงคุณภาพ หรือข้อมูลนับ

กราฟแท่ง

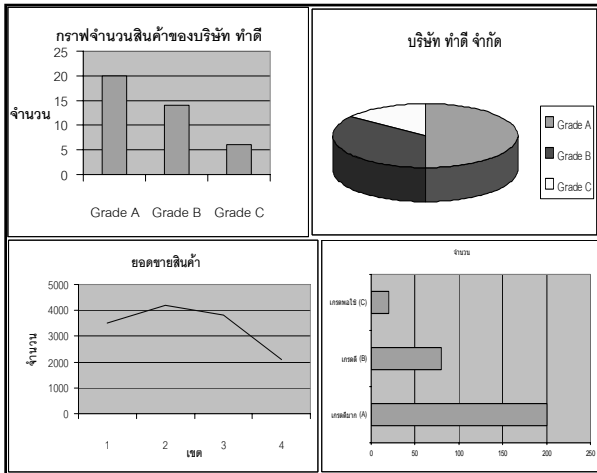
29

ตารางแสดงจำนวนสินค้าที่ผลิตได้จำแนกตามคุณภาพของสินค้า

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

ข้อมูลจำแนกประเภท

30




การแจกแจงความถี่

1. การแจกแจงแบบไม่จัดกลุ่ม

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



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

33

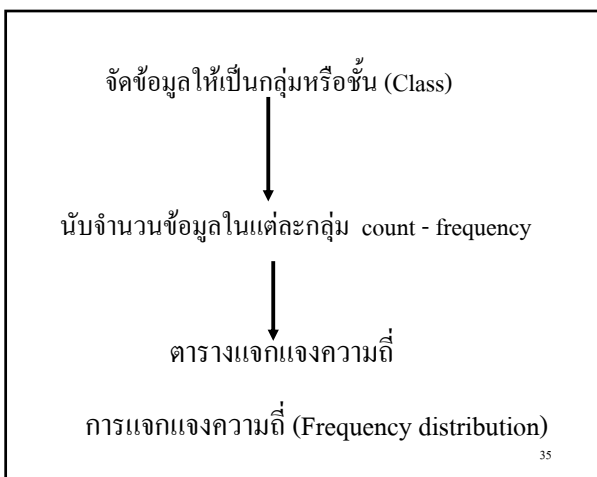
2. การแจกแจงแบบจัดกลุ่ม

ตัวอย่าง อายุ (ปี) ของลูกค้าของห้างสรรพสินค้าแห่งหนึ่งจำนวน 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

34



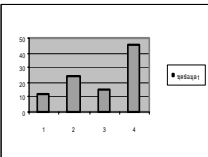
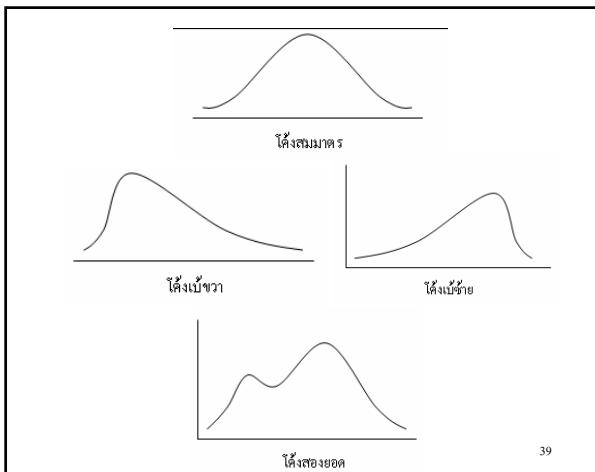
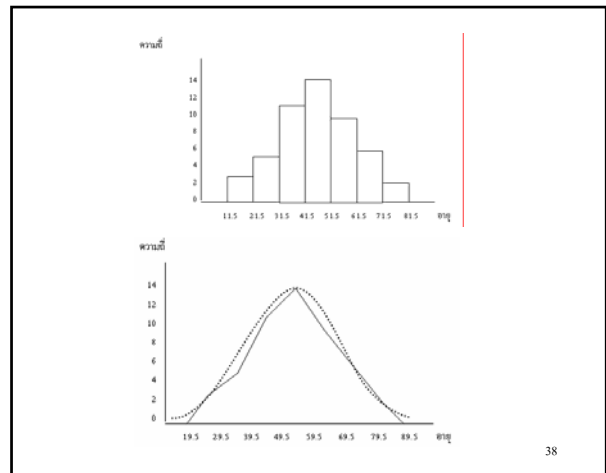
อายุ (ปี) (ขีดจำกัดชั้น)	รอยคะแนน	ความถี่
12 - 21		3
22 - 31		5
32 - 41		11
42 - 51		14
52 - 61		9
62 - 71		6
72 - 81		2
รวม		50

36

นำค่าข้อมูลและความถี่ของแต่ละกลุ่มสร้างกราฟ

↓

แผนภาพการแจกแจงความถี่ :
แผนภูมิแท่ง (Histogram)
ฮิสโตแกรม

๑ ตำแหน่งข้อมูล

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

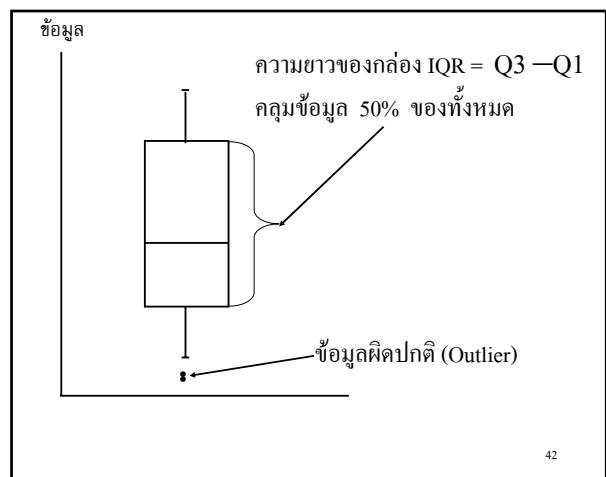
ควอไทล์ เป็นค่าข้อมูลที่อยู่ในตำแหน่งต่างๆ เมื่อแบ่งข้อมูลเป็น 4 ส่วนเท่าๆ กัน เรียกค่าที่อยู่ในแต่ละตำแหน่งนั้นว่า ควอไทล์ที่ 1 ควอไทล์ที่ 2 และ ควอไทล์ที่ 3 หรือคือ Q_1, Q_2 และ Q_3

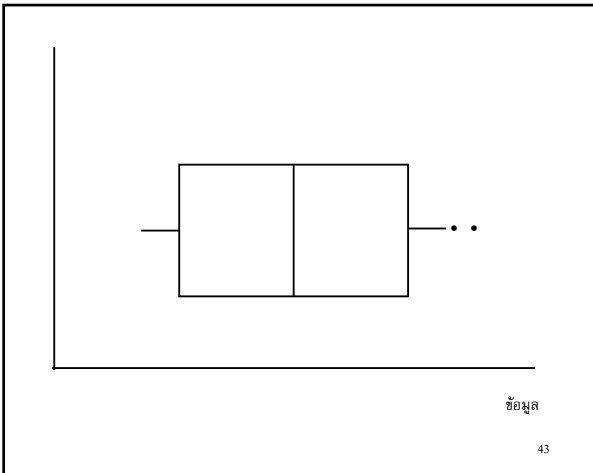
ซึ่ง $Q_1 = P_{25}$, $Q_2 = P_{50}$ และ $Q_3 = P_{75}$

แผนภาพกล่อง (Box and Whisker plot)

เป็นแผนภาพที่ใช้สรุปลักษณะของข้อมูล แผนภาพกล่องจะแสดงลักษณะที่สำคัญของข้อมูลชุดนั้นๆ เช่น ค่ากลาง ค่าการกระจาย ลักษณะการแจกแจงข้อมูล และค่าผิดปกติของข้อมูล (outlier)

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





แผนภาพลำต้นและใบ Stem and leaf plot

เป็นแผนภาพแสดงการแจกแจงความถี่ประกอบด้วย 2 ส่วนคือ ส่วน Stem และส่วน leaf

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

๓ ค่ากลาง

เพื่ออธิบายลักษณะข้อมูลด้วยค่าตัวเลข

ตัววัดแนวโน้มสู่ส่วนกลางหรือค่ากลางที่สำคัญ คือ

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

ค่าเฉลี่ย (Average หรือ Mean)

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 ข้อมูลต่อไปนี้เป็นระยะทางวิ่งต่อน้ำมัน 1 ลิตรของรถขี้อ่อนหนึ่ง 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})

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

ค่ามัธยฐาน (median)

ค่ามัธยฐาน (median) เป็นค่ากลางที่แบ่งข้อมูลเป็น 2 ส่วนเท่าๆ กันจะมีข้อมูลจำนวนครึ่งหนึ่งที่มีค่าต่ำกว่ามัธยฐาน และอีกครึ่งหนึ่งที่มีค่าสูงกว่าค่ามัธยฐาน $\text{ค่ามัธยฐาน} = \frac{X_{n+1}}{2}$

การหาค่ามัธยฐาน เรียงข้อมูลจำนวน n ค่าจากน้อยไปมาก
- ถ้า n เป็นจำนวนคี่ ค่ามัธยฐานคือ ค่าข้อมูลที่อยู่ตำแหน่งกึ่งกลาง
- ถ้า n เป็นจำนวนคู่ ค่ามัธยฐานคือ ค่าเฉลี่ยระหว่างข้อมูล 2 ค่าที่อยู่ตำแหน่งกึ่งกลาง

49

ตัวอย่าง ข้อมูล 2 ชุดต่อไปนี้คือ

ชุด A : 11 13 18 19 20

ชุด B : 4 4 9 11 12 14

ค่ามัธยฐานของข้อมูลชุด A คือ 18

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

50

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

การใช้ค่าฐานนิยมจะใช้ได้ทั้งในข้อมูลเชิงปริมาณและเชิงคุณภาพ

51

ตัวอย่าง การใช้ฐานนิยมในข้อมูลเชิงคุณภาพ

เกรด B เป็นฐานนิยมของวิชาหลัก สถิติ 1

สมชาย เป็นชื่อที่เป็นฐานนิยมของชายไทย

เลือกกลุ่ม O เป็นฐานนิยมของกลุ่มเลือดของคน

52

4 ค่าวัดการกระจายหรือความผันแปร (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 ดังนั้นควรใช้ทั้งค่ากลางและค่าวัดการกระจายในการอธิบายข้อมูล

53

ค่าสถิติที่ใช้วัดการกระจาย หรือความผันแปร

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

54

❶ พิสัย range

พิสัย = ข้อมูลค่าสูงสุด - ค่าต่ำสุด

ค่าพิสัยมาก แสดงว่าข้อมูลกระจายหรือผันแปรมาก

❷ ค่าพิสัยระหว่างควอไทล์ Interquartile range

$$IQR = Q_3 - Q_1$$

IQR วัดการกระจายได้ดีกว่าค่าพิสัย เพราะเมื่อมีค่าสูงหรือต่ำเกินไป จะมีผลต่อ IQR น้อยกว่าพิสัย

55

❸ ค่าความแปรปรวนและ ❹ ค่าเบี่ยงเบนมาตรฐาน

เป็นค่าวัดความผันแปรหรือวัดการกระจายของข้อมูลที่นิยมใช้กันมาก เนื่องจากเป็นค่าที่ใช้ข้อมูลทุกค่ามาคำนวณ

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

56

❸.2 การประมวลผลเชิงอนุมาน

เป็นการจัดการข้อมูลด้วยสถิติเชิงอนุมานเพื่ออธิบายข้อมูลทั้งหมดด้วยข้อมูลตัวอย่าง ประกอบด้วย การประมาณค่า และการทดสอบสมมติฐาน

เป็นการคำนวณและวิเคราะห์ข้อมูลตัวอย่าง

57

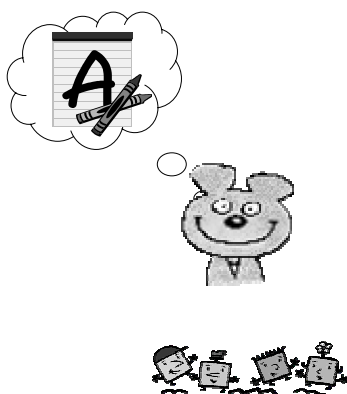
การประมาณค่า เป็นวิธีการคำนวณค่าข้อมูลตัวอย่างเพื่อตอบค่าพารามิเตอร์ซึ่งเป็นค่าอธิบายลักษณะประชากร

ค่าพารามิเตอร์มีค่าเท่าไร ?

การทดสอบสมมติฐาน เป็นวิธีการที่นำข้อมูลตัวอย่างมาสรุปไปสู่ประชากรเพื่อทดสอบว่าพารามิเตอร์เป็นไปตามที่สงสัยหรือคาดการณ์ไว้หรือไม่

พารามิเตอร์มีค่าตามที่คาดไว้หรือไม่

58



59