

สถิติและฐานข้อมูลสำหรับนักบริหาร

(Statistics and database for administrators)



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

สถิติและฐานข้อมูลสำหรับนักบริหาร

- ข้อมูลสถิติสำหรับการบริหาร
- กระบวนการดำเนินงานทางสถิติ
- ข้อมูลสถิติ การเก็บรวบรวมข้อมูล คลังข้อมูล
- การอธิบายข้อมูล
- การวัดแนวโน้มสู่ค่ากลางและการกระจายข้อมูล
- การแจกแจงปกติ

17-18-24-25 ธ.ค. 2554 และ 14-15 ม.ค. 2555



ความหมายของสถิติ (Statistics : Statistik : State)

- สถิติ : **ตัวเลข**ที่ใช้บรรยายเหตุการณ์หรือข้อเท็จจริง (facts) ของเรื่องต่างๆ ที่เราต้องการศึกษา สามารถนำมา**เปรียบเทียบ วิเคราะห์หาความหมาย**ได้
 - สถิติ : **ศาสตร์หรือวิชา**ที่ว่าด้วยการเก็บรวบรวมข้อมูล การวิเคราะห์ข้อมูล การแปลความหมายข้อมูล และการนำเสนอข้อมูล เพื่อนำผลการวิเคราะห์ที่ได้มาช่วยประกอบการตัดสินใจ ซึ่งสถิติในความหมายนี้ มักเรียกว่า สถิติศาสตร์ (Statistics)
- สถิติ : การเดาอย่างมีระบบ ?**



สถิติและฐานข้อมูลสำหรับนักบริหาร

กระบวนการดำเนินงานทางสถิติ

1. การวางแผน (planning)
2. การเก็บรวบรวมข้อมูล (data collection)
3. การนำเสนอข้อมูล (data presentation)
4. การวิเคราะห์ข้อมูล (data analysis)
5. การแปลความหมาย (data interpretation)



การใช้สถิติสำหรับนักบริหาร



การวางแผน (planning)

Organization

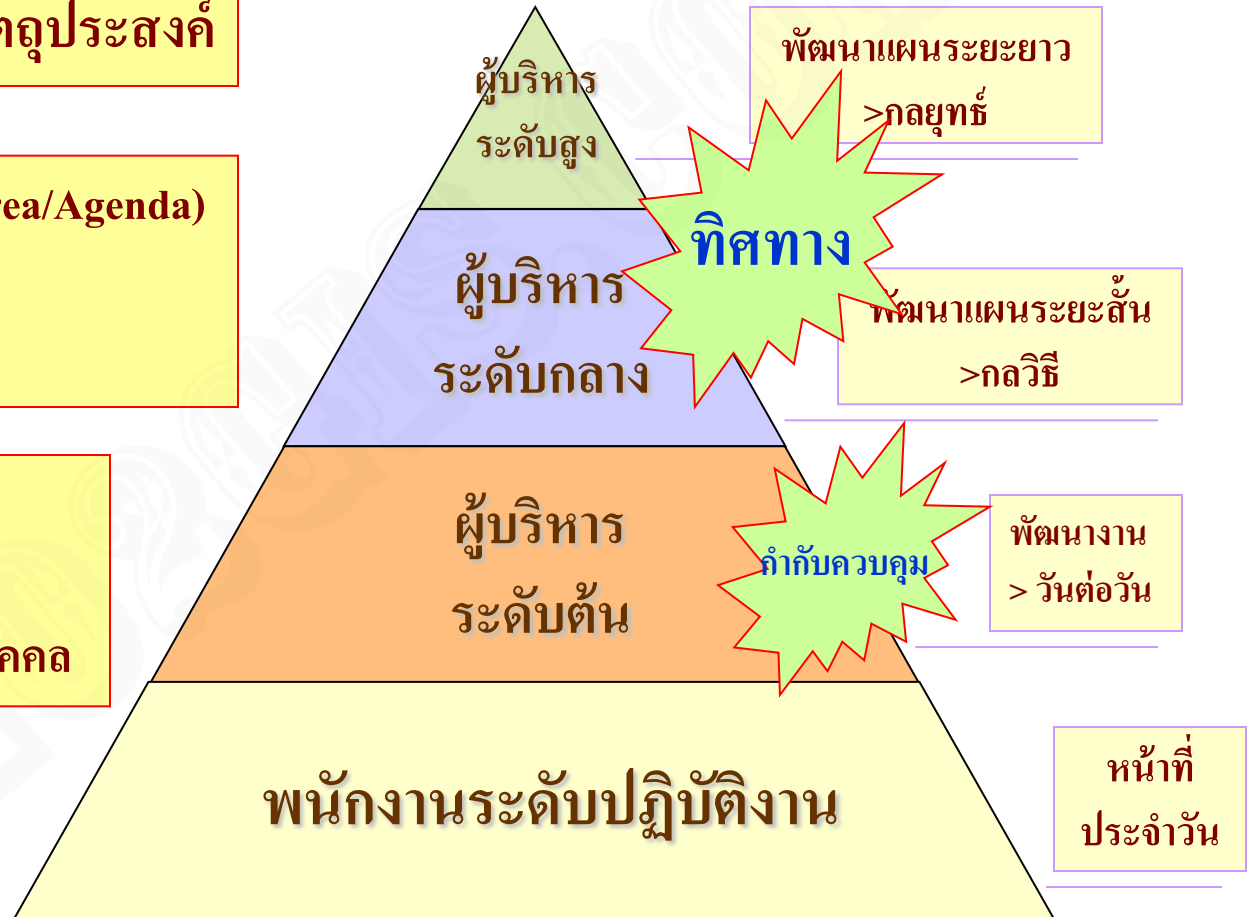


วัตถุประสงค์

ภารกิจองค์การ (Functions /Area/Agenda)
แนวคิด
กระบวนการ

Management

1. คน / งาน / เงิน / ทรัพยากร
2. สาทกล / ประเทศ / องค์การ / บุคคล



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

การเก็บรวบรวมข้อมูล (data collection)



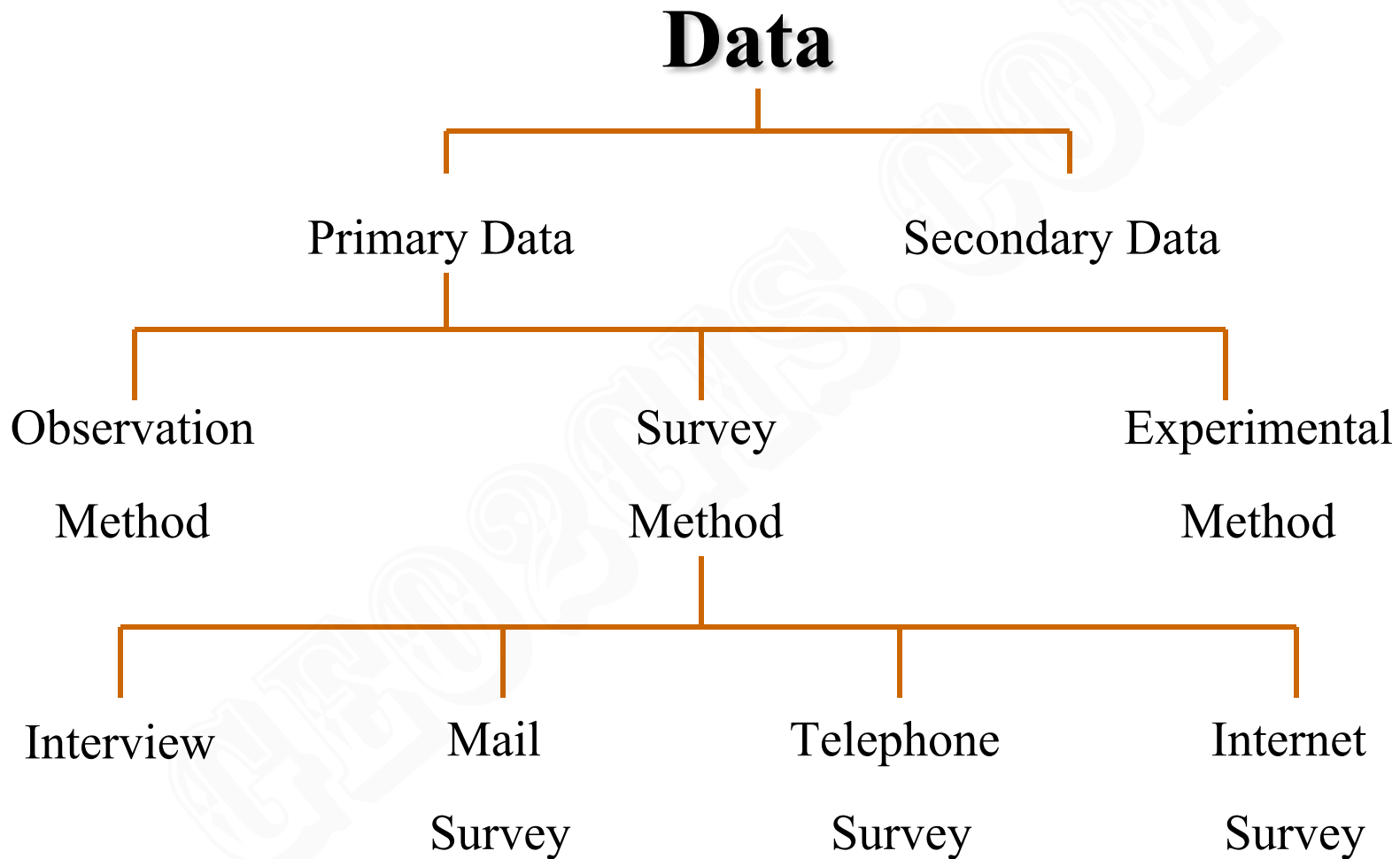
สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

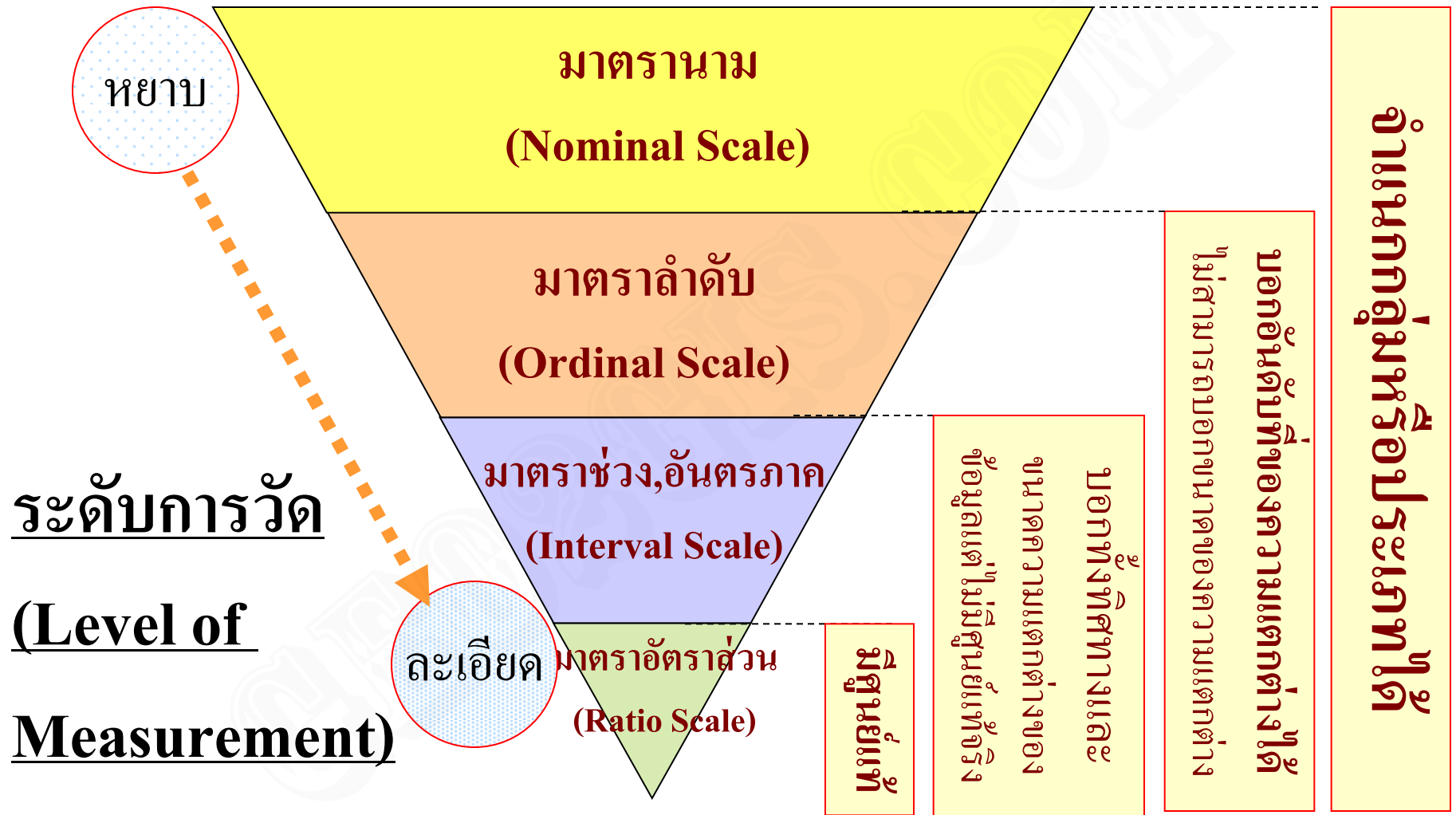
การใช้สถิติสำหรับ**นักบริหาร** ➡ **การเก็บรวบรวมข้อมูล (data collection)**



การเก็บรวบรวมข้อมูล (data collection)



ข้อมูลสถิติ



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

นายอนุสร พุ่มพวง

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

anu.pumpuang@gmail.com

การใช้สถิติสำหรับนักบริหาร



การเก็บรวบรวมข้อมูล

การสุ่มตัวอย่างข้อมูล

ประชากร



ความน่าจะเป็น > ขนาดเหมาะสม

กลุ่มตัวอย่าง



ค่าพารามิเตอร์

เช่น μ

ค่าสถิติ

เช่น $\bar{X}$

ประมาณค่าเพื่ออ้างอิงประชากร



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

การใช้สถิติสำหรับนักบริหาร ➡ การเก็บรวบรวมข้อมูล

การกำหนดขนาดของกลุ่มตัวอย่าง

% ความคลาดเคลื่อน

100

50

0

100

200

จำนวนกลุ่มตัวอย่าง



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

การใช้สถิติสำหรับนักบริหาร ➡ การเก็บรวบรวมข้อมูล

วิธีการสุ่มตัวอย่าง

ใช้หลักความน่าจะเป็น

1. การสุ่มตัวอย่างแบบง่าย
2. การสุ่มตัวอย่างแบบมีระบบ
3. การสุ่มตัวอย่างแบบแบ่งชั้น
4. การสุ่มตัวอย่างแบบกลุ่ม
5. การสุ่มตัวอย่างหลายขั้นตอน



การใช้สถิติสำหรับนักบริหาร ➡ การเก็บรวบรวมข้อมูล

วิธีการสุ่มตัวอย่าง

ไม่ใช่หลักความน่าจะเป็น

1. การสุ่มตัวอย่างแบบบังเอิญ หรือการสุ่มตัวอย่างตามความสะดวก
2. การสุ่มตัวอย่างแบบเจาะจง
3. การสุ่มตัวอย่างแบบโควต้า
4. การสุ่มตัวอย่างโดยใช้วิธีบอกต่อ



สรุปการใช้สถิติสำหรับนักบริหาร

สถิติเชิงพรรณนา (Descriptive Statistics) บรรยายลักษณะข้อมูล

ที่รวบรวมมา โดยสรุปลักษณะสำคัญ ประกอบด้วย

- การนำเสนอข้อมูล
- การนำเสนอตัวแทนของข้อมูล
- การนำเสนอตำแหน่งหรือลำดับของข้อมูล
- การอธิบายการกระจายของข้อมูล
- การอธิบายการแจกแจงความถี่ของข้อมูล



ตารางแจกแจงความถี่

สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

การแจกแจงความถี่

รายการ	รอยขีด	ความถี่
4	//	2
5	/	2
9	//	2
12	//	2
16	///	3
17	///	3
19	////	4
รวม	18	18

19 4 4 16 17 19 17 9 9
5 12 12 19 5 16 16 17 19

รายการ	รอยขีด	ความถี่
4 – 7	////	4
8 – 11	//	2
12 – 15	//	2
16 – 19	////////	10
รวม	18	18

ไม่จัดกลุ่มข้อมูล

จัดกลุ่มข้อมูล



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

การสร้างตารางแจกแจงความถี่

19 5 5 16 17 19 17 9 9 5 12 12 19 5 16 16 17 19

1. หาผลต่างระหว่างค่าสูงสุด – ต่ำสุด (ค่าพิสัย :

$$19 - 5$$

Range : R)

2. กำหนดจำนวนชั้น (Class : C)

3

3. หาค่าความห่างของข้อมูลแต่ละชั้น (อันตรภาคชั้น

: Class Interval : I)

$$I = \frac{R}{C}$$

$$\frac{19 - 5}{3} = 4.6 = 5$$

4. เขียนขีดจำกัดชั้นข้อมูลทั้งหมด

(โดยปกติใช้ข้อมูลค่าน้อยที่สุดเป็นขีดจำกัดล่างของชั้นที่ 1)

รายการ	ความถี่
5 – 9	6
10 – 14	2
15 – 19	10
รวม	18



สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

การแจกแจงความถี่สะสม

19 5 5 16 17 19 17 9 9 5 12 12 19 5 16 16 17 19

อายุราชการ	ความถี่	ความถี่สะสม
5 – 7	4	4
8 – 10	2	6
11 – 13	2	8
14 – 16	3	11
17 – 19	7	18

- ความถี่ 7 หมายความว่า
มีข้าราชการอยู่ 7 คนที่มีอายุราชการ
อยู่ระหว่าง 16.5 ปี และ 19.5 ปี
- ความถี่สะสม 8 หมายความว่า
มีข้าราชการอยู่ 8 คนที่มีอายุราชการ
ต่ำกว่า 13.5 ปี



สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

การแจกแจงความถี่สัมพัทธ์

19 5 5 16 17 19 17 9 9 5 12 12 19 5 16 16 17 19

อายุราชการ	ความถี่	ความถี่สัมพัทธ์	ร้อยละของ ความถี่สัมพัทธ์
5 – 7	4	$4/18 = 0.22$	22
8 – 10	2	$2 / 18 = 0.11$	11
11 – 13	2	$2 / 18 = 0.11$	11
14 – 16	3	$3 / 18 = 0.17$	17
17 – 19	7	$7 / 18 = 0.39$	39
รวม	18	1.00	100

- ร้อยละของความถี่สัมพัทธ์ 17
หมายความว่า มีข้าราชการอยู่
17 % ที่มีอายุราชการ
13.5 ปี ถึง 16.5 ปี



การนำเสนอข้อมูล

การนำเสนอข้อมูล



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

จำนวนผู้ป่วยโรคเอดส์ในเขตกรุงเทพมหานครตั้งแต่ พ.ศ. 2527- 2550

	ทั้งหมด	เสียชีวิต	ยังมีชีวิตอยู่
2527 - 2544	19,508	5,892	13,616
2545	2,669	656	2,013
2546	2,741	585	2,156
2547	2,909	466	2,443
2548	3,048	348	2,700
2549	3,689	470	3,219
2550	2,287	303	1,984
รวม	36,851	8,720	28,131

แหล่งข้อมูล : ฝ่ายแผนงานและระบาดวิทยา กองควบคุม
โรคเอดส์ สำนักอนามัย กรุงเทพมหานคร

ร้อยละ (Percentage)

: การคำนวณค่าสัดส่วนจากฐาน 100

ร้อยละ =

ค่าตัวเลขที่ต้องการคำนวณค่าร้อยละ x 100

ค่าตัวเลขทั้งหมด

1. ร้อยละของจำนวนผู้เสียชีวิตที่เป็นผู้ป่วยโรคเอดส์
ใน กทม. ปี พ.ศ. 2550

$$= \frac{303 \times 100}{2,287}$$

$$= 13.25$$

2. ร้อยละของจำนวนผู้ที่ยังมีชีวิตอยู่ที่เป็นผู้ป่วยโรค
เอดส์ใน กทม. ปี พ.ศ. 2550

$$= \frac{1,984 \times 100}{2,287}$$

$$= 86.25$$



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

จำนวนผู้ป่วยโรคเอดส์ในเขตกรุงเทพมหานครตั้งแต่ พ.ศ. 2527- 2550

	ทั้งหมด	เสียชีวิต	ยังมีชีวิตอยู่
2527 - 2544	19,508	5,892	13,616
2545	2,669	656	2,013
2546	2,741	585	2,156
2547	2,909	466	2,443
2548	3,048	348	2,700
2549	3,689	470	3,219
2550	2,287	303	1,984
รวม	36,851	8,720	28,131

แหล่งข้อมูล : ฝ่ายแผนงานและระบาดวิทยา กองควบคุม
โรคเอดส์ สำนักอนามัย กรุงเทพมหานคร

อัตราส่วน (Ratio)

: การเปรียบเทียบระหว่างค่าสองค่า

$$\text{อัตราส่วน} = \frac{\text{ค่าตัวเลขที่ 1}}{\text{ค่าตัวเลขที่ 2}}$$

อัตราส่วนระหว่างจำนวนผู้เสียชีวิต
กับจำนวนผู้ที่ยังมีชีวิตอยู่ ที่เป็นผู้ป่วย
โรคเอดส์ใน กทม. ปี พ.ศ. 2550

$$= \frac{303}{1,984}$$

$$= 1 : 6.5$$



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

จำนวนผู้ป่วยโรคเอดส์ในเขตกรุงเทพมหานครตั้งแต่ พ.ศ. 2527- 2550

สัดส่วน (Proportion)

	ทั้งหมด	เสียชีวิต	ยังมีชีวิตอยู่
2527 - 2544	19,508	5,892	13,616
2545	2,669	656	2,013
2546	2,741	585	2,156
2547	2,909	466	2,443
2548	3,048	348	2,700
2549	3,689	470	3,219
2550	2,287	303	1,984
รวม	36,851	8,720	28,131

: การเปรียบเทียบระหว่างค่าตัวเลขน้อย
กับค่าตัวเลขรวมทั้งหมด

$$\text{สัดส่วน} = \frac{\text{ค่าตัวเลขน้อย}}{\text{ค่าตัวเลขทั้งหมด}}$$

ผู้ป่วยโรคเอดส์ใน กทม. ปี พ.ศ. 2550

$$\begin{aligned} \text{สัดส่วนผู้เสียชีวิต} &= \frac{303}{2,287} \\ &= 1 : 7.5 \end{aligned}$$

$$\begin{aligned} \text{สัดส่วนผู้ยังมีชีวิต} &= \frac{1,984}{2,287} \\ &= 1 : 1.15 \end{aligned}$$

แหล่งข้อมูล : ฝ่ายแผนงานและระบาดวิทยา กองควบคุม
โรคเอดส์ สำนักอนามัย กรุงเทพมหานคร



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

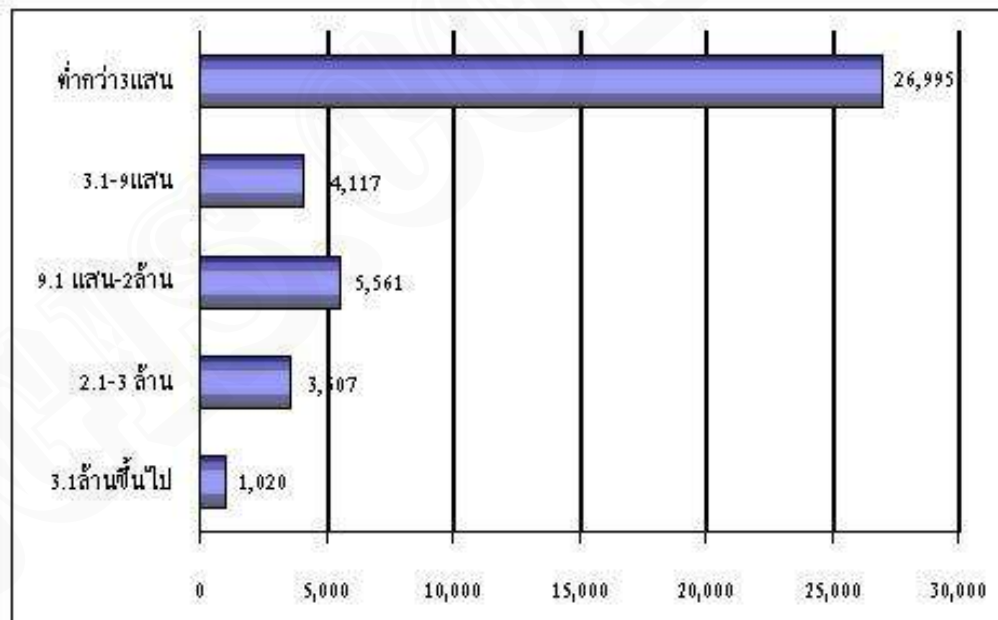
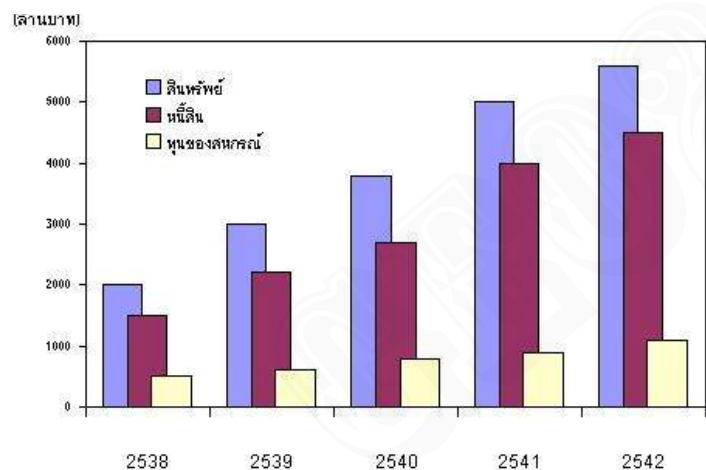
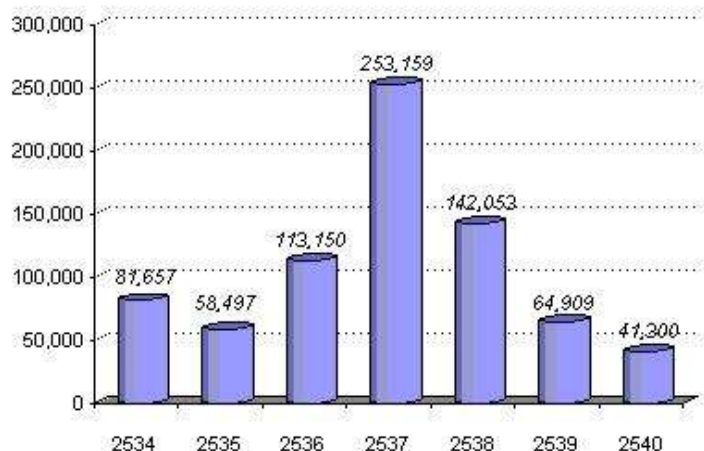
นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

การใช้สถิติสำหรับนักบริหาร



การนำเสนอข้อมูล



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

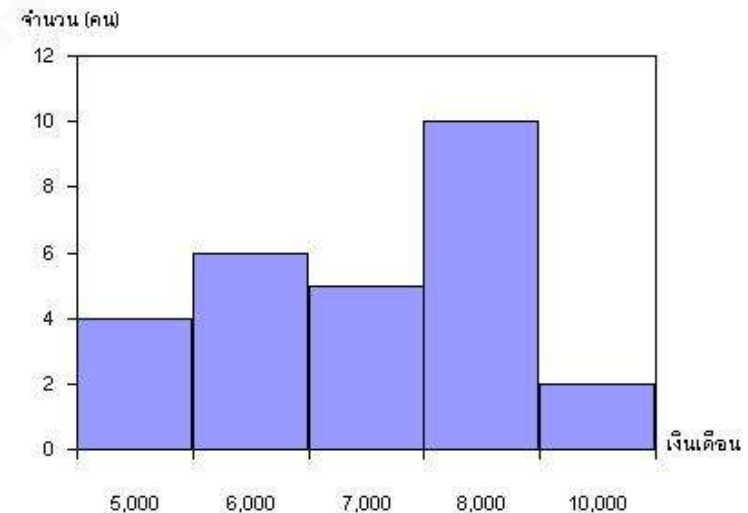
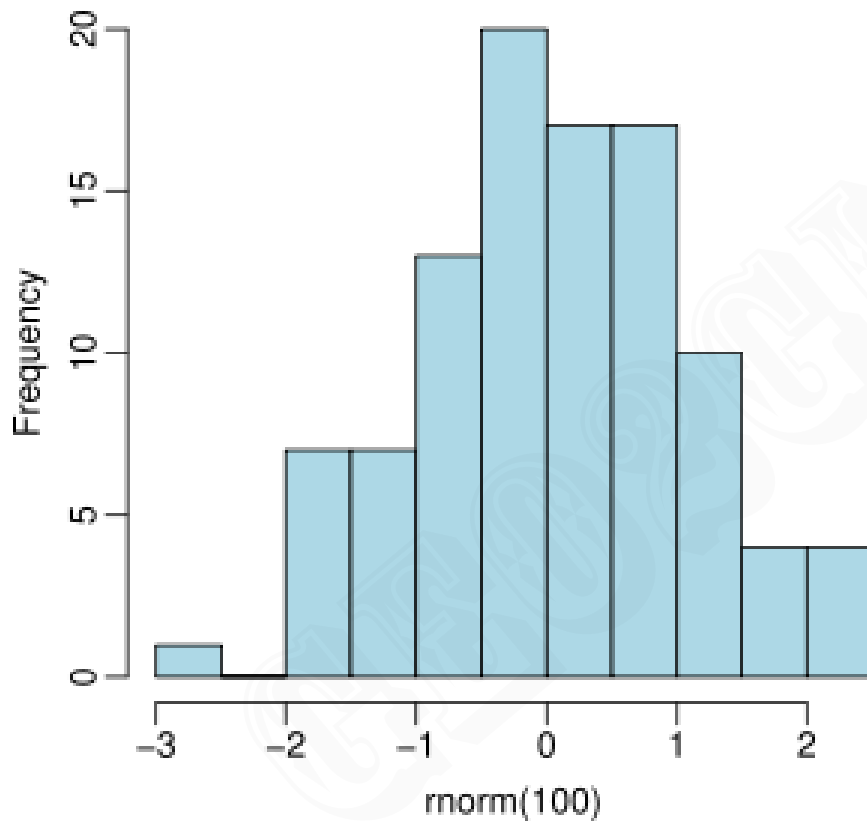
วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

การใช้สถิติสำหรับนักบริหาร ➡ การนำเสนอข้อมูล

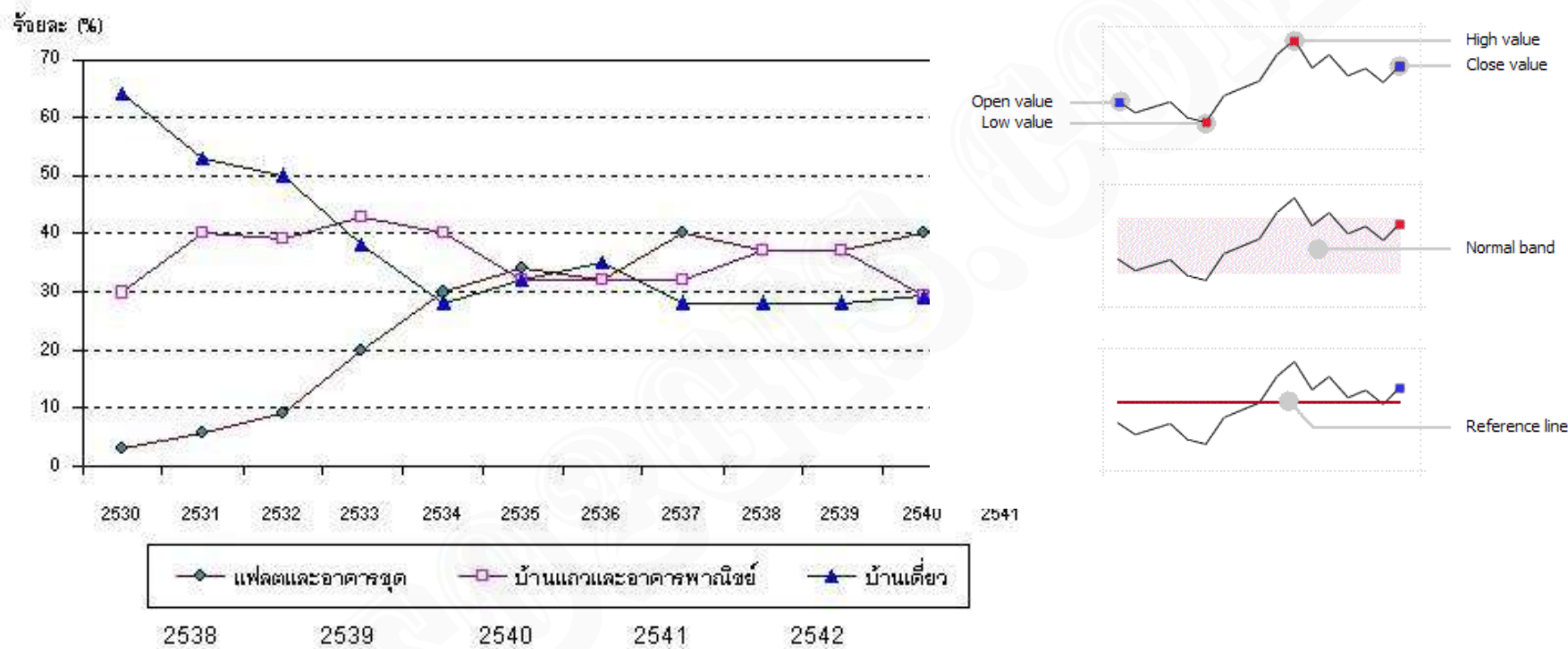
Histogram of rnorm(100)



การใช้สถิติสำหรับนักบริหาร



การนำเสนอข้อมูล



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

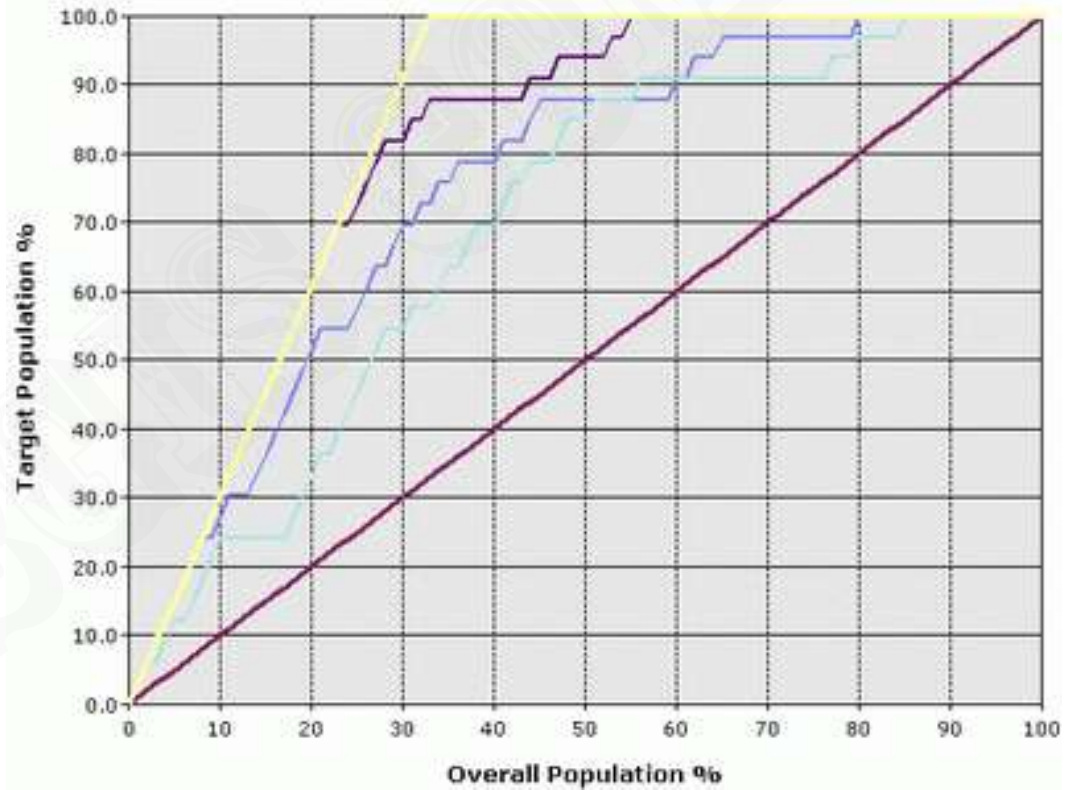
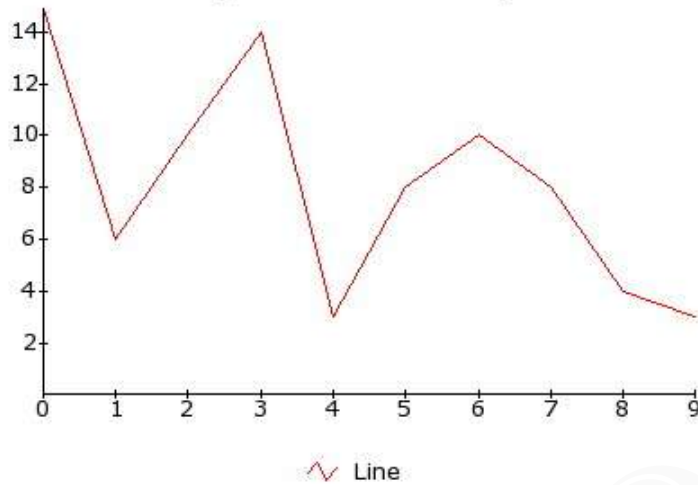
anu.pumpuang@gmail.com

การใช้สถิติสำหรับ**นักบริหาร**



การนำเสนอข้อมูล

Simple Line Chart Sample



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

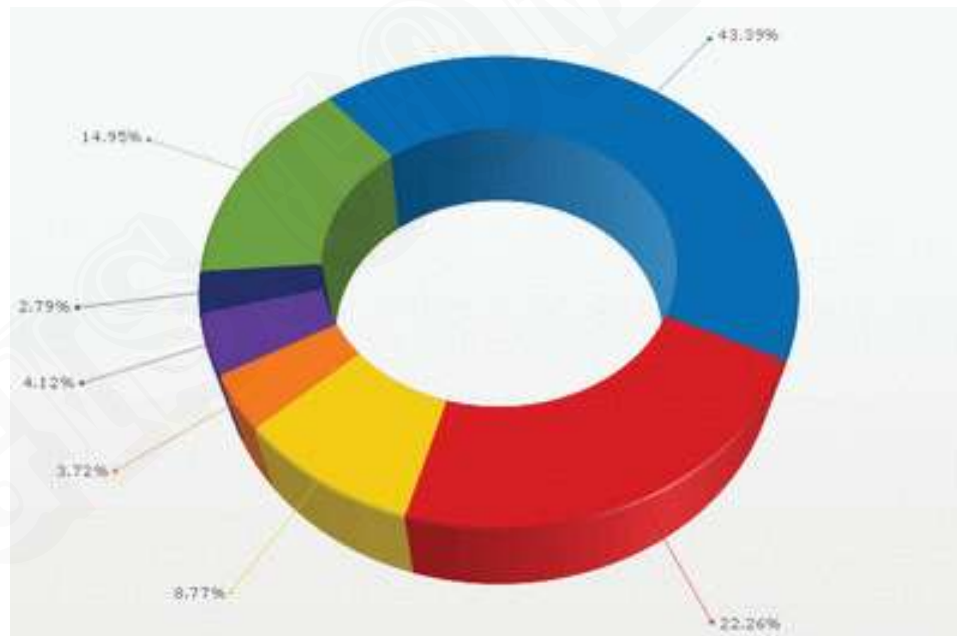
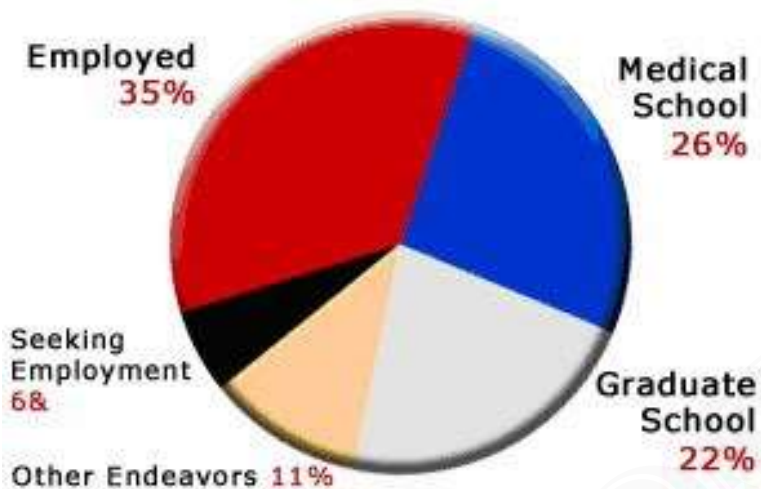
นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

การใช้สถิติสำหรับนักบริหาร



การนำเสนอข้อมูล



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

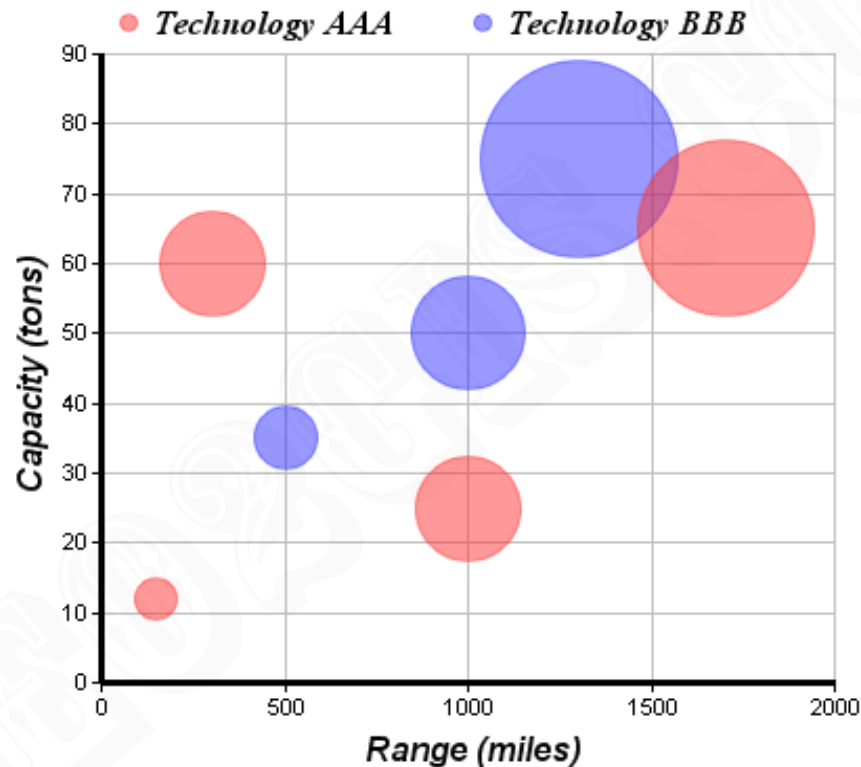
วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

การใช้สถิติสำหรับ นักบริหาร ➡ การนำเสนอข้อมูล

Product Comparison Chart



การวัดแนวโน้มสู่ค่ากลาง

การหาค่ากลางของข้อมูลเพื่อใช้เป็นตัวแทนของ
ข้อมูลทั้งหมด เพื่อเป็นประโยชน์ในการเปรียบเทียบ
ข้อมูลแต่ละชุด โดยไม่จำเป็นต้องพิจารณาข้อมูล
ทั้งหมดของแต่ละชุด

การวัดแนวโน้มสู่ค่ากลาง

ค่าเฉลี่ย

(Mean)



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

การวัดแนวโน้มสู่ค่ากลาง

ค่าเฉลี่ยเลขคณิต (Arithmetic Mean)

สำหรับข้อมูลที่ไม่แจกแจงความถี่

19 5 5 16 17 19 17 9 9 5 12 12 19 5 16 16 17 19

$$\bar{X} = \frac{\sum X}{N}$$

$$\bar{X} = \frac{19+5+5+16+17+19+17+9+9+5+12+12+19+5+16+16+17+19}{18}$$

237/18

13.17



การวัดแนวโน้มสู่ค่ากลาง

ค่าเฉลี่ยเลขคณิต (Arithmetic Mean)

สำหรับข้อมูลที่แจกแจงความถี่

19 5 5 16 17 19 17 9 9 5 12 12 19 5 16 16 17 19

อายุราชการ (x)	ความถี่ (f)	fx
5	4	20
9	2	18
12	2	24
16	3	48
17	3	51
19	4	76
รวม	18	237

$$\bar{X} = \frac{\sum fx}{N}$$

237/18

13.17



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

การวัดแนวโน้มสู่ค่ากลาง

ค่าเฉลี่ยเลขคณิต (Arithmetic Mean)

สำหรับข้อมูลที่แจกแจงความถี่

19 5 5 16 17 19 17 9 9 5 12 12 19 5 16 16 17 19

อายุราชการ	ความถี่ (f)	จุดกลาง (x)	fx
5 – 7	4	6	24
8 – 10	2	9	18
11 – 13	2	12	24
14 – 16	3	15	45
17 – 19	7	18	126
	N = 18		$\Sigma fx = 237$

$$\bar{X} = \frac{\Sigma fx}{N}$$

237/18

13.17



การวัดแนวโน้มสู่ค่ากลาง

ค่ามัธยฐาน (Median)



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

การวัดแนวโน้มสู่ค่ากลาง

ค่ามัธยฐาน (Median) :

ข้อมูลที่อยู่ตรงกลางของข้อมูลทั้งหมด เมื่อได้เรียงลำดับข้อมูลเหล่านั้นแล้ว

สำหรับข้อมูลที่ไม่แจกแจงความถี่

ข้อมูลเป็นจำนวนคี่

5 5 5 5 9 9 12 12 16 16 16 17 17 17 19 19 19

ค่ามัธยฐาน (Median) = 16 ปี

ข้อมูลเป็นจำนวนคู่

5 5 5 5 9 9 12 12 16 16 16 17 17 17 19 19 19 19

ค่ามัธยฐาน (Median) = $\frac{16 + 16}{2} = 16$ ปี



การวัดแนวโน้มสู่ค่ากลาง

ค่ามัธยฐาน (Median) : สำหรับข้อมูลที่แจกแจงความถี่

$$Mdn = L_0 + i \left[\frac{\frac{N}{2} - cf}{f} \right]$$

หาชั้นที่ตำแหน่งของค่ามัธยฐานอยู่

$$\frac{N}{2} = 18/2 = 9$$

ชั้นที่ค่ามัธยฐานอยู่คือ 14 – 16

$$L_0 = 13.5$$

$$i = 3$$

$$cf = 8$$

$$f = 3$$

$$Mdn = 13.5 + 3 \left[\frac{9 - 8}{3} \right] = 14.5 \text{ ปี}$$

อายุราชการ	ความถี่ (f)	ความถี่สะสม
5 – 7	4	4
8 – 10	2	6
11 – 13	2	8
14 – 16	3	11
17 – 19	7	18



การวัดแนวโน้มสู่ค่ากลาง

! ค่าฐานนิยม (Mode)



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

การวัดแนวโน้มสู่ค่ากลาง

ค่าฐานนิยม (Mode) : **คะแนนที่มีความถี่มากที่สุด**
สำหรับข้อมูลที่ไม่แจกแจงความถี่

17 19 17 18 18 19 17 16 17

ฐานนิยม = 17

17 19 17 18 19 19 16 16 17

ฐานนิยม = 17 และ 19

17 15 14 18 13 11 16 12 19

ข้อมูลชุดนี้ไม่มีฐานนิยม



การวัดแนวโน้มสู่ค่ากลาง

ค่าฐานนิยม (Mode) : คะแนนที่มีความถี่มากที่สุด
สำหรับข้อมูลที่แจกแจงความถี่

$$Mo = L_0 + i \left[\frac{\triangle_1}{\triangle_1 + \triangle_2} \right]$$

$$\triangle_1 = 7 - 2 = 5$$

$$\triangle_2 = 7 - 3 = 4$$

$$Mo = 10.5 + 3 \left[\frac{5}{5 + 4} \right]$$

$$= 12.16$$

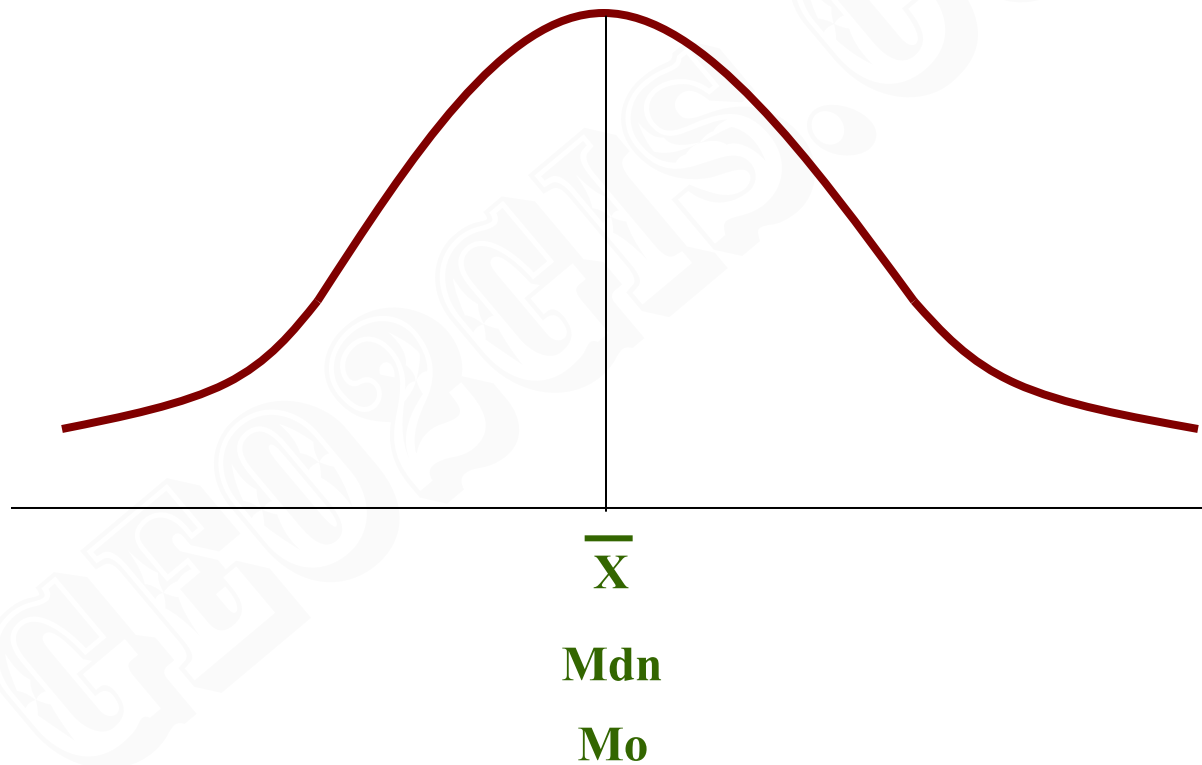
อายุราชการ	ความถี่ (f)	
5 – 7	4	
8 – 10	2	f0
11 – 13	7	f1
14 – 16	3	f2
17 – 19	5	



การวัดแนวโน้มสู่ค่ากลาง

ความสัมพันธ์ระหว่างค่าเฉลี่ยเลขคณิต มัธยฐาน และฐานนิยม

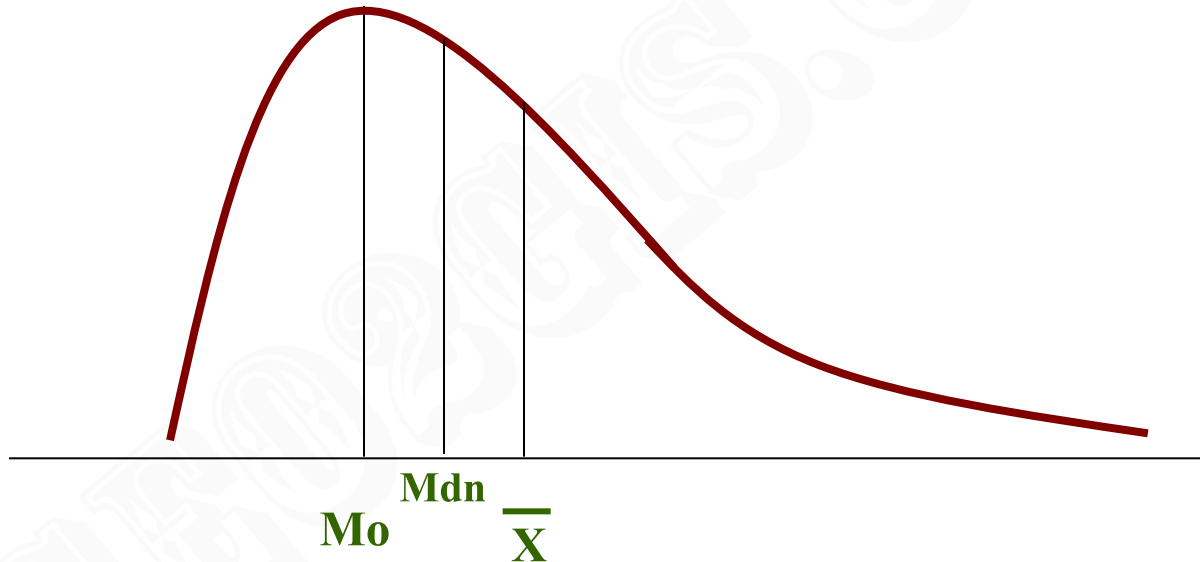
การแจกแจงแบบปกติ



การวัดแนวโน้มสู่ค่ากลาง

ความสัมพันธ์ระหว่างค่าเฉลี่ยเลขคณิต มัธยฐาน และฐานนิยม

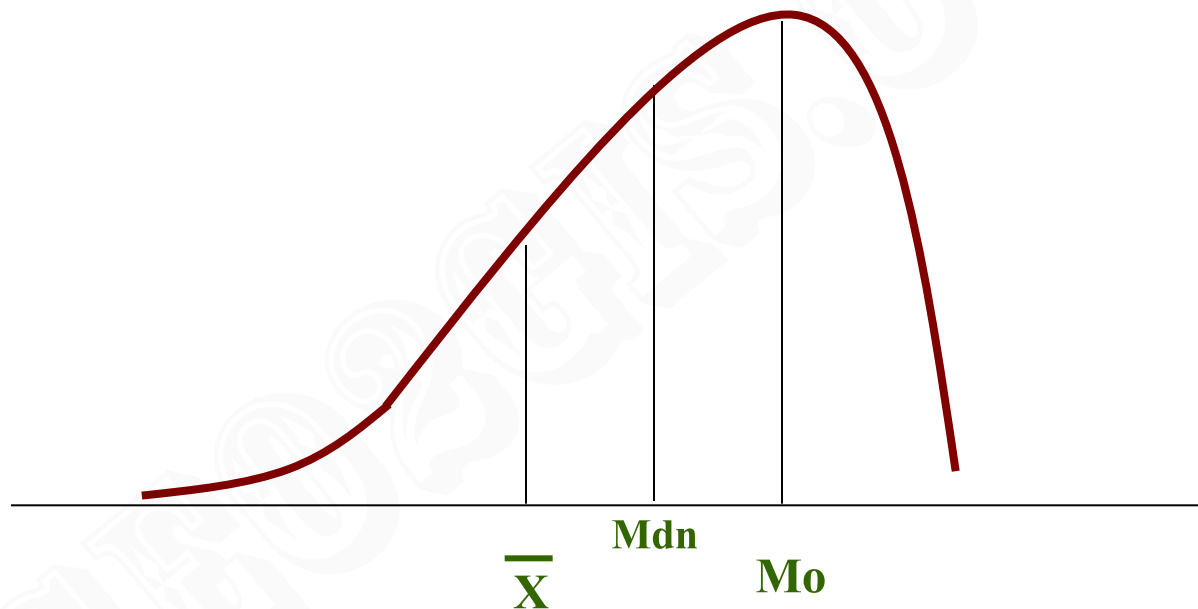
การแจกแจงเบ้ทางบวก



การวัดแนวโน้มสู่ค่ากลาง

ความสัมพันธ์ระหว่างค่าเฉลี่ยเลขคณิต มัธยฐาน และฐานนิยม

การแจกแจงเบ้ทางลบ



การวัดแนวโน้มสู่ค่ากลาง

การเลือกใช้แนวโน้มเข้าสู่ส่วนกลาง

มาตราการวัด	แนวโน้มเข้าสู่ส่วนกลางที่ใช้
มาตรานาม (Nominal Scale)	ฐานนิยม
มาตราลำดับ (Ordinal Scale)	มัธยฐานหรือฐานนิยม
มาตราช่วง, อันตรภาค (Interval Scale)	ค่าเฉลี่ยเลขคณิต มัธยฐานหรือฐานนิยม
มาตราอัตราส่วน (Ratio Scale)	ค่าเฉลี่ยเลขคณิต มัธยฐานหรือฐานนิยม



การหาค่าตำแหน่งข้อมูล



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

การหาค่าตำแหน่งข้อมูล

1. มัธยฐาน (Median : Mdn)

เป็นค่าที่แบ่งข้อมูลออกเป็น 2 ส่วนเท่าๆกัน

2. ควอไทล์ (Quartile : Q)

เป็นค่าที่แบ่งข้อมูลออกเป็น 4 ส่วนเท่าๆกัน

3. เดซิอัล (Decile : D)

เป็นค่าที่แบ่งข้อมูลออกเป็น 10 ส่วนเท่าๆกัน

4. เพอร์เซนต์ไทล์ (Percentile : P)

เป็นค่าที่แบ่งข้อมูลออกเป็น 100 ส่วนเท่าๆกัน



ค่าควอไทล์ (Quatile)



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

ค่าควอไทล์ (Quatile)

สำหรับข้อมูลที่ไม่แจกแจงความถี่

$$Q_r = \text{ตำแหน่งที่ } (N + 1) \frac{r}{4}$$

จากข้อมูล 12 11 22 25 22 25 11 15 ให้หาค่า **Quatile ที่ 2**

1. เรียงลำดับใหม่ 11 11 12 15 22 22 25 25

2. หา $Q_2 = \text{ตำแหน่งที่ } (8 + 1) \frac{2}{4} = 4.5$

3. ตำแหน่งที่ 4.5 คือค่ากึ่งกลางระหว่างตำแหน่งที่ 4 และตำแหน่งที่ 5

ซึ่งมีค่าระหว่าง 15 กับ 22 จึงมีค่า = $\frac{15 + 22}{2} = 27$



ค่าควอไทล์ (Quatile) : สำหรับข้อมูลที่แจกแจงความถี่

จากข้อมูล หาค่าควอไทล์ (Quatile) ที่ 3 ตำแหน่งของข้อมูล = $\frac{rN}{4}$

คะแนน	ความถี่ (f)	ความถี่สะสม
1-9	4	4
10-19	6	10
20-29	8	18
30-39	10	28
40-49	4	32 = Q_3
50-59	6	38
60-69	2	40
รวม	40	

$$= \frac{3 \times 40}{4} = 30$$

$$Q_r = L_0 + i \left[\frac{\frac{rN}{4} - cf}{f} \right]$$

$$Q_3 = 39.5 + 10 \left[\frac{30 - 28}{4} \right] \\ = 44.5$$



ค่าเฉลี่ย (Decile)



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

ค่าเฉลี่ย (Decile)

สำหรับข้อมูลที่ไม่แจกแจงความถี่

$$D_r = \text{ตำแหน่งที่ } \frac{(N + 1) \cdot r}{10}$$

จากข้อมูล 12 11 22 25 22 25 11 15 16 18 22 12 13 14 ให้หาค่า **Decile ที่ 5**

1. เรียงลำดับใหม่ 11 11 12 **12 13 14 15 16 18** 22 22

2. หา $D_5 = \text{ตำแหน่งที่ } \frac{(14 + 1) \cdot 5}{10} = 7.5$

3. ตำแหน่งที่ 7.5 คือค่ากึ่งกลางระหว่างตำแหน่งที่ 7 และตำแหน่งที่ 8

ซึ่งมีค่าระหว่าง 15 กับ 22 จึงมีค่า = $\frac{15 + 16}{2} = 15.5$



ค่าเฉลี่ย (Decile) : สำหรับข้อมูลที่แจกแจงความถี่

จากข้อมูล หาค่าเฉลี่ย (Decile) ที่ 3

คะแนน	ความถี่ (f)	ความถี่สะสม
1-9	4	4
10-19	6	10
20-29	8	18 = D_3
30-39	10	28
40-49	4	32
50-59	6	38
60-69	2	40
รวม	40	

$$\text{ตำแหน่งของข้อมูล} = \frac{rN}{10}$$

$$= \frac{3 \times 40}{10} = 12$$

$$D_r = L_0 + i \left[\frac{\frac{rN}{10} - cf}{f} \right]$$

$$D_3 = 19.5 + 10 \left[\frac{12 - 10}{8} \right]$$
$$= 22$$



ค่าเปอร์เซ็นต์ไทล์ (Percentile)



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

ค่าเปอร์เซ็นต์ไทล์ (Percentile)

สำหรับข้อมูลที่ไม่แจกแจงความถี่

$$P_r = \text{ตำแหน่งที่ } \frac{(N + 1) \cdot r}{100}$$

จากข้อมูล 12 11 22 25 22 25 11 15 ให้หาค่า **Percentile ที่ 50**

1. เรียงลำดับใหม่ 11 11 12 15 22 22 25 25

2. หา $P_{50} = \text{ตำแหน่งที่ } \frac{(8 + 1) \cdot 50}{100} = 4.5$

3. ตำแหน่งที่ 4.5 คือค่าระหว่างตำแหน่งที่ 4 และตำแหน่งที่ 5 ซึ่งมีค่าระหว่าง 15 กับ 22

3. ตำแหน่งที่ 5 มากกว่าตำแหน่งที่ 4 = 1 ลำดับ มีข้อมูลต่างกันเท่ากับ $22 - 15 = 8$

ตำแหน่งที่ 4.5 มากกว่าตำแหน่งที่ 4 = 0.5 ลำดับ มีข้อมูลต่างกันเท่ากับ $\frac{8 \times 0.5}{1} = 4$

ดังนั้น ตำแหน่งที่ 4.5 จึงมีค่าข้อมูลเท่ากับ $15 + 4 = 19$



ค่าเปอร์เซ็นต์ไทล์ (Percentile) : สำหรับข้อมูลที่แจกแจงความถี่

จากข้อมูล หาค่าเปอร์เซ็นต์ไทล์ (Percentile) ที่ 70

คะแนน	ความถี่ (f)	ความถี่สะสม
1-9	4	4
10-19	6	10
20-29	8	18
30-39	10	28 = D ₃
40-49	4	32
50-59	6	38
60-69	2	40
รวม	40	

$$\begin{aligned}\text{ตำแหน่งของข้อมูล} &= \frac{rN}{100} \\ &= \frac{70 \times 40}{100} = 28\end{aligned}$$

$$D_r = L_0 + i \left[\frac{\frac{rN}{10} - cf}{f} \right]$$

$$\begin{aligned}D_3 &= 29.5 + 10 \left[\frac{28 - 18}{10} \right] \\ &= 39.5\end{aligned}$$



การอธิบายการกระจายของข้อมูล

การกระจายข้อมูล

การวัดการกระจายของข้อมูลเป็นการอธิบายว่า
ข้อมูลแต่ละค่านั้นมีค่าที่ห่างกันมากน้อยเพียงใด

ค่าการกระจายของข้อมูล**มาก** แสดงว่าข้อมูลชุดนั้น
มีค่าแตกต่างกันมาก

ค่าการกระจายของข้อมูล**น้อย** แสดงว่าข้อมูลชุดนั้น
มีค่าซ้ำ ๆ กัน ค่าใกล้เคียงกัน



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

การกระจายข้อมูล

พิสัย

(Range)



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

พิสัย (Range) :

ขีดจำกัดบนของคะแนนสูงสุด - ขีดจำกัดล่างของคะแนนต่ำสุด

ค่าสูงสุด - ค่าต่ำสุด

ก : 1 2 4 5 6 8 10

ข : 1 1 8 9 9 10 10

$$R = 10 - 1 = 9$$

คะแนน	ความถี่ (f)
1-9	4
10-19	6
20-29	8
30-39	10

$$R = 39.5 - 0.5 = 39$$

ควรหลีกเลี่ยงวัดการกระจายข้อมูลโดยใช้ค่าพิสัย
สำหรับข้อมูลมีขนาดใหญ่ หรือข้อมูลจำนวนมาก



การกระจายข้อมูล

ส่วนเบี่ยงเบนควอไทล์ (Quatile Deviation)



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

ส่วนเบี่ยงเบนควอไทล์ (Quatile Deviation) : QD

การวัดการกระจายข้อมูลโดยใช้ค่าส่วนเบี่ยงเบนมาตรฐาน (SD)

จะใช้ในการนิข้อมูลชุดนั้นวัดแนวโน้มเข้าสู่ส่วนกลางโดยใช้

ค่ามัธยฐาน (Mdn)

$$QD = \frac{Q3 - Q1}{2}$$



การกระจายข้อมูล

ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) : SD

การวัดการกระจายข้อมูลโดยใช้ค่าส่วนเบี่ยงเบนมาตรฐาน (SD)

จะใช้ในกรณีข้อมูลชุดนั้นวัดแนวโน้มเข้าสู่ส่วนกลางโดยใช้

ค่าเฉลี่ยเลขคณิต ($\bar{X}$, μ)

ส่วนเบี่ยงเบนมาตรฐาน SD : ค่าสถิติ (ใช้กับกลุ่มตัวอย่าง)

σ (Sigma) : ค่าพารามิเตอร์ (ใช้กับประชากร)

ความแปรปรวน SD^2 : ค่าสถิติ (ใช้กับกลุ่มตัวอย่าง)

σ^2 : ค่าพารามิเตอร์ (ใช้กับประชากร)



ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) : SD

สำหรับข้อมูลที่ไม่แจกแจงความถี่

สำหรับข้อมูลที่แจกแจงความถี่

ประชากร

$$\sigma = \sqrt{\frac{\sum (x - \mu)^2}{N}}$$

SD : ส่วนเบี่ยงเบนมาตรฐาน

X : ค่าคะแนนแต่ละจำนวน

σ : ค่าเฉลี่ยของข้อมูลชุดนั้น

N : จำนวนข้อมูลของชุดนั้น

f : ค่าความถี่

$$\sigma = \sqrt{\frac{\sum f(x - \mu)^2}{N}}$$

$$\sigma = \sqrt{\frac{\sum fx^2}{N} - \left(\frac{\sum fx}{N}\right)^2}$$

$$\sigma = i \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2}$$



ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) : SD

สำหรับข้อมูลที่ไม่แจกแจงความถี่

$$SD = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

SD : ส่วนเบี่ยงเบนมาตรฐาน

X : ค่าคะแนนแต่ละจำนวน

X : ค่าเฉลี่ยของข้อมูลชุดนั้น

n : จำนวนข้อมูลของชุดนั้น

f : ค่าความถี่

สำหรับข้อมูลที่แจกแจงความถี่

$$SD = \sqrt{\frac{\sum f(x - \bar{x})^2}{n - 1}}$$

กลุ่มตัวอย่าง



การกระจายข้อมูล

ส่วนเบี่ยงเบนเฉลี่ย (Mean Deviation)



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

ส่วนเบี่ยงเบนเฉลี่ย (Mean Deviation) : MD

ค่าเฉลี่ยของส่วนเบี่ยงเบนระหว่างข้อมูลแต่ละจำนวน กับ
ค่าเฉลี่ยเลขคณิตของข้อมูลชุดนั้น

สำหรับข้อมูลที่ไม่แจกแจงความถี่

$$MD = \frac{\sum |(X - \bar{X})|}{N}$$

สำหรับข้อมูลที่แจกแจงความถี่

$$MD = \frac{\sum f |(X - \bar{X})|}{N}$$

MD : ส่วนเบี่ยงเบนเฉลี่ย

X : ข้อมูลแต่ละจำนวน

$\bar{X}$: ค่าเฉลี่ยเลขคณิต

f : ความถี่

N : จำนวนข้อมูลทั้งหมด

|| : ค่าสัมบูรณ์ (ตัวเลขภายใน
ไม่คิดเครื่องหมาย)



การกระจายข้อมูล

สัมประสิทธิ์ความแปรผัน (Coefficient of Variance)



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

สัมประสิทธิ์ความแปรผัน (Coefficient of Variance)

เหมาะสำหรับการวัดการกระจายข้อมูลของข้อมูล 2 ชุด

- เปรียบเทียบการกระจายข้อมูล 2 ชุดที่มีค่าเฉลี่ยเลขคณิต หรือค่ามัธยฐาน ต่างกันมาก ๆ

- เปรียบเทียบการกระจายข้อมูล 2 ชุดที่มีหน่วยต่างกัน เช่น กก. / ปอนด์

น้ำหนักของเด็ก นร. ที่สมัครเรียนเทควันโด

ห้อง A : 35 35 40 45 45 SD = 25

ห้อง B : 20 25 25 60 70 SD = 53.75

น้ำหนักเฉลี่ยของเด็ก นร. ทั้ง 2 ห้องเท่ากับ 40 กิโลกรัม



สัมประสิทธิ์ความแปรผัน (Coefficient of Variance)

เมื่อวัดการกระจายด้วยพิสัย

$$CV = \frac{\text{ค่าสูงสุด} - \text{ค่าต่ำสุด}}{\text{ค่าสูงสุด} + \text{ค่าต่ำสุด}} \times 100 \%$$

เมื่อวัดการกระจายด้วยส่วนเบี่ยงเบนควอไทล์

$$CV = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100 \%$$

เมื่อวัดการกระจายด้วยส่วนเบี่ยงเบนเฉลี่ย

$$CV = \frac{MD}{\bar{X}} \times 100 \%$$

เมื่อวัดการกระจายด้วยส่วนเบี่ยงเบนมาตรฐาน

$$CV = \frac{SD}{\bar{X}} \times 100 \%$$

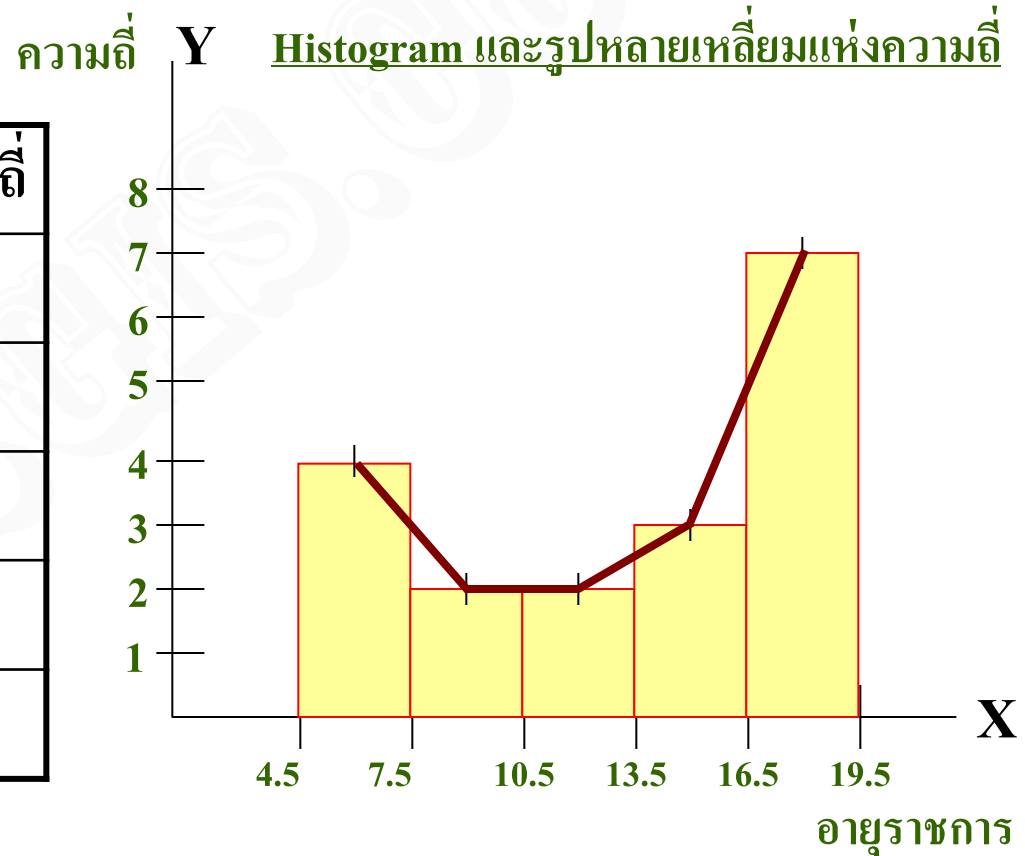
CV ต่ำ : ข้อมูลชุดนั้นกระจายน้อย

CV สูง : ข้อมูลชุดนั้นกระจายมาก



การอธิบายการแจกแจงความถี่ของข้อมูล

อายุราชการ	ขีดจำกัดชั้นจริง	ความถี่
5 – 7	4.5 - 7.5	4
8 – 10	7.5 - 10.5	2
11 – 13	10.5 - 13.5	2
14 – 16	13.5 - 16.5	3
17 – 19	16.5 - 19.5	7

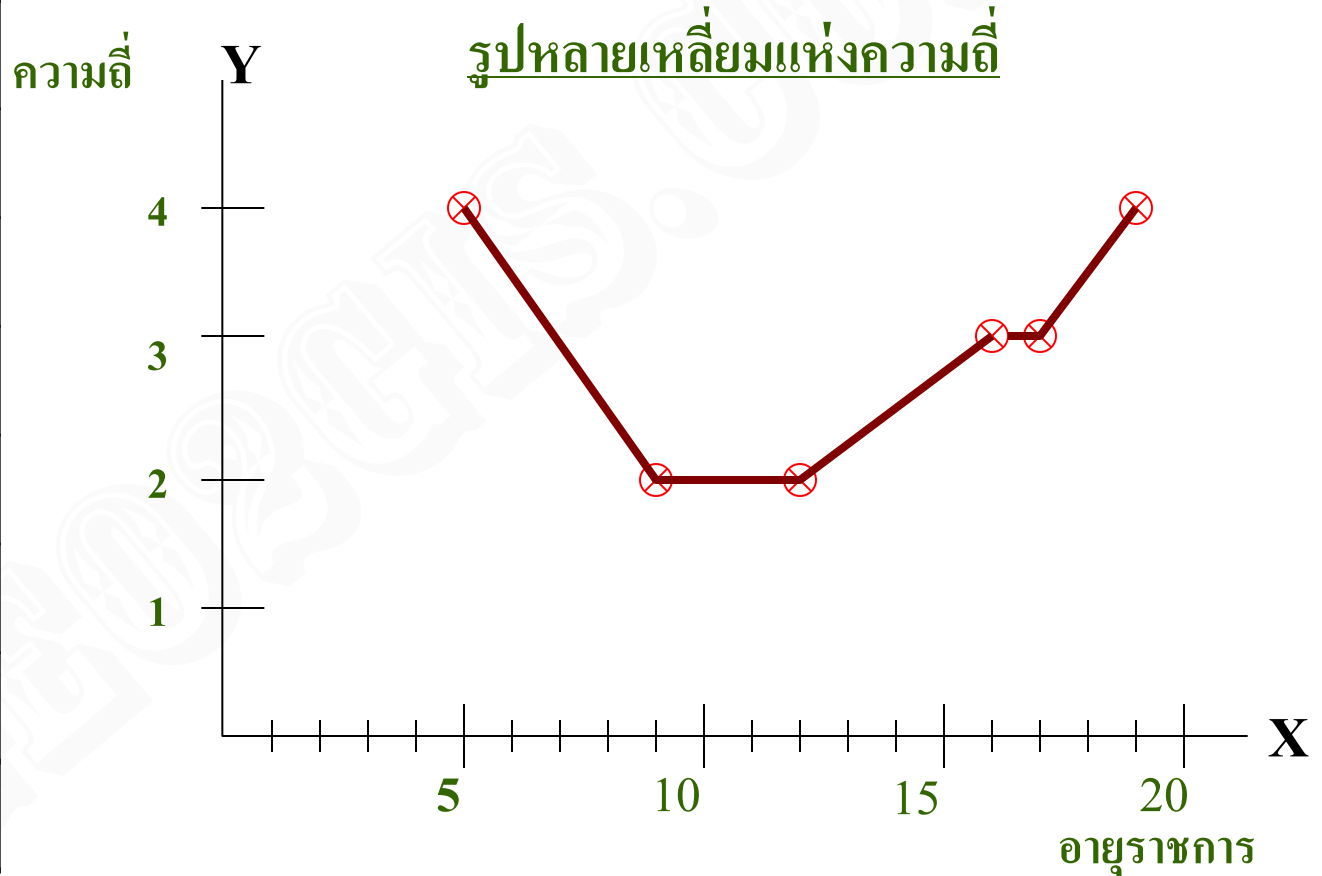


สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

การแจกแจงความถี่ด้วยแผนภูมิหรือกราฟ

19 5 5 16 17 19 17 9 9 5 12 12 19 5 16 16 17 19

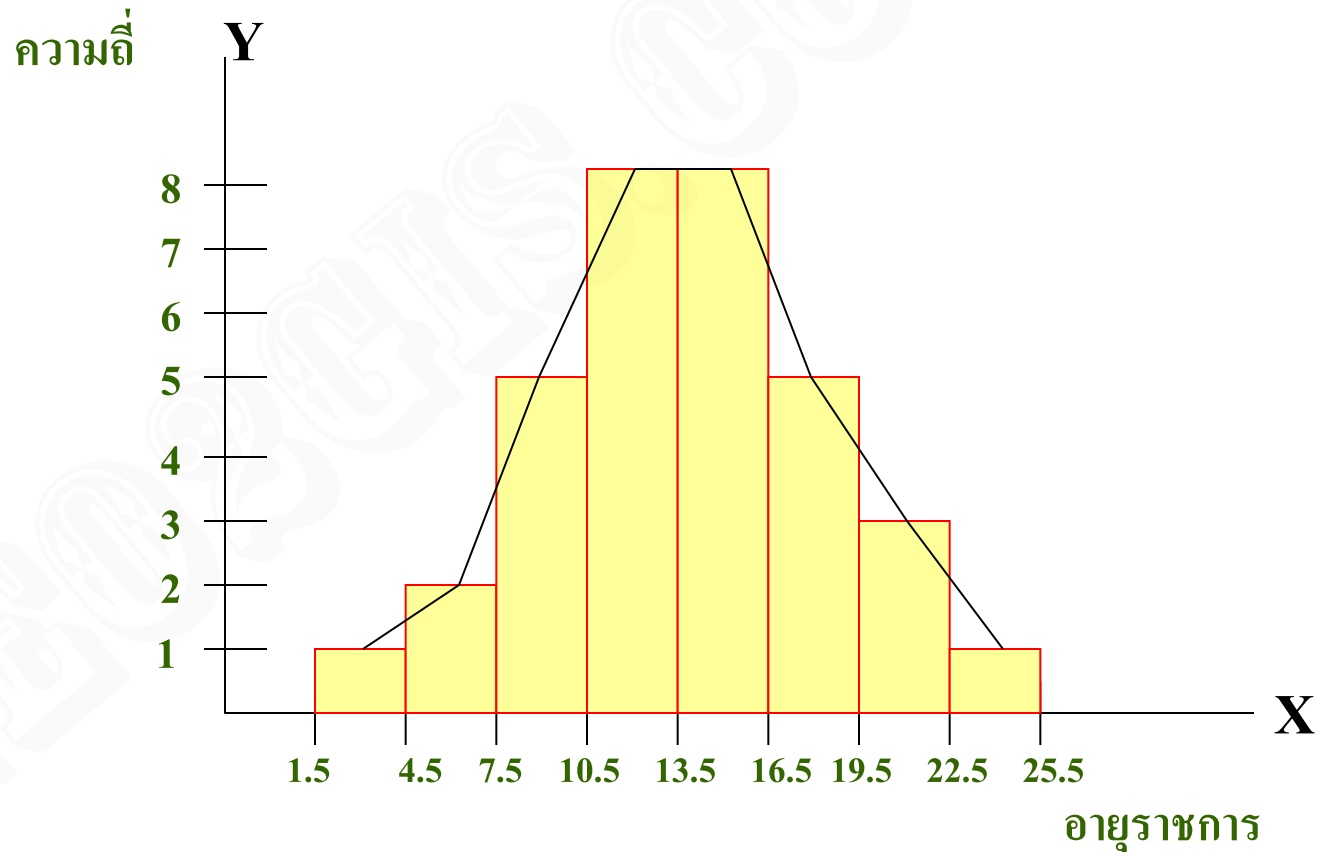
อายุราชการ	ความถี่
5	4
9	2
12	2
16	3
17	3
19	4
รวม	18



สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

การแจกแจงความถี่ด้วยแผนภูมิหรือกราฟ

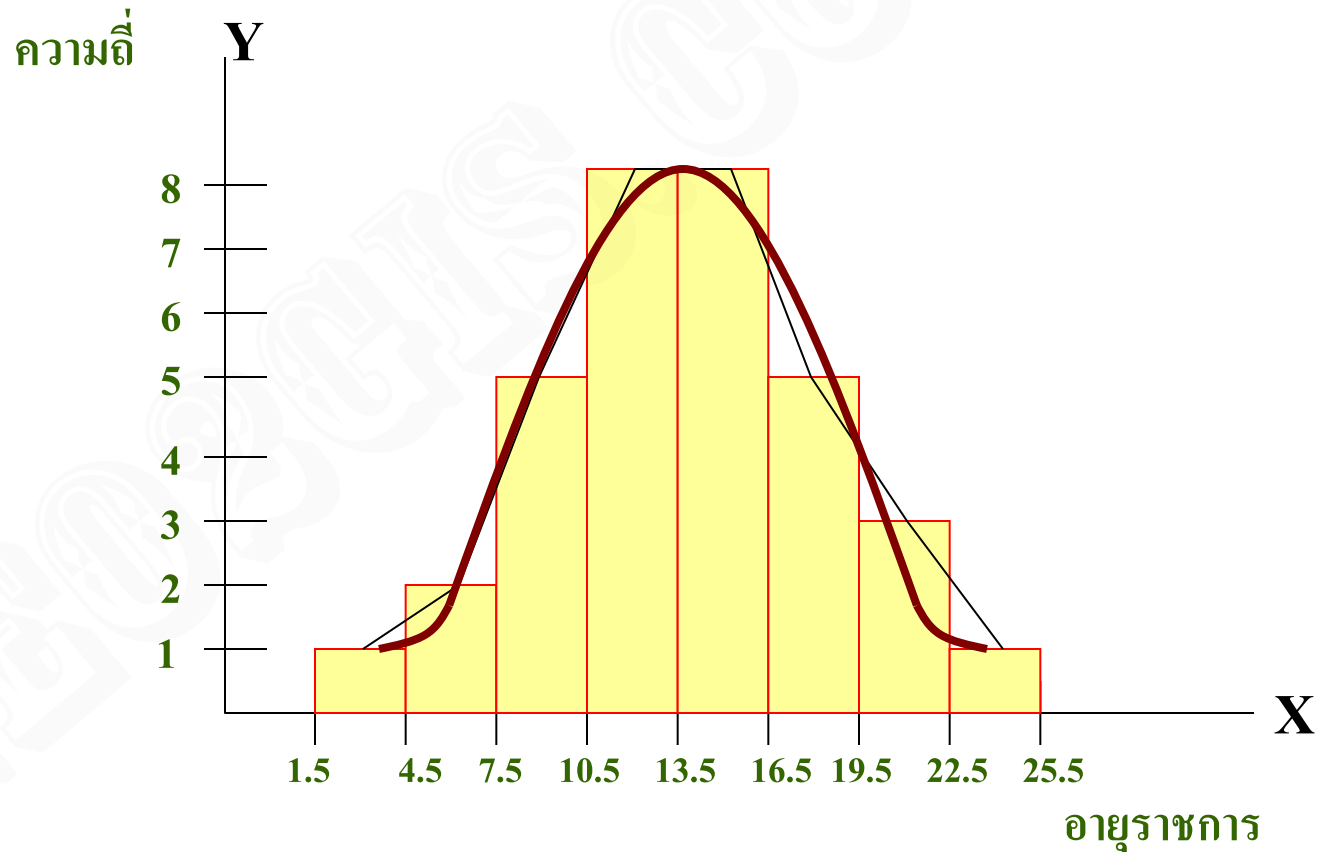
อายุราชการ	ความถี่
2 - 4	1
5 - 7	2
8 - 10	5
11 - 13	8
14 - 16	8
17 - 19	5
20 - 22	3
23 - 25	1



สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

การแจกแจงความถี่ด้วยแผนภูมิหรือกราฟ

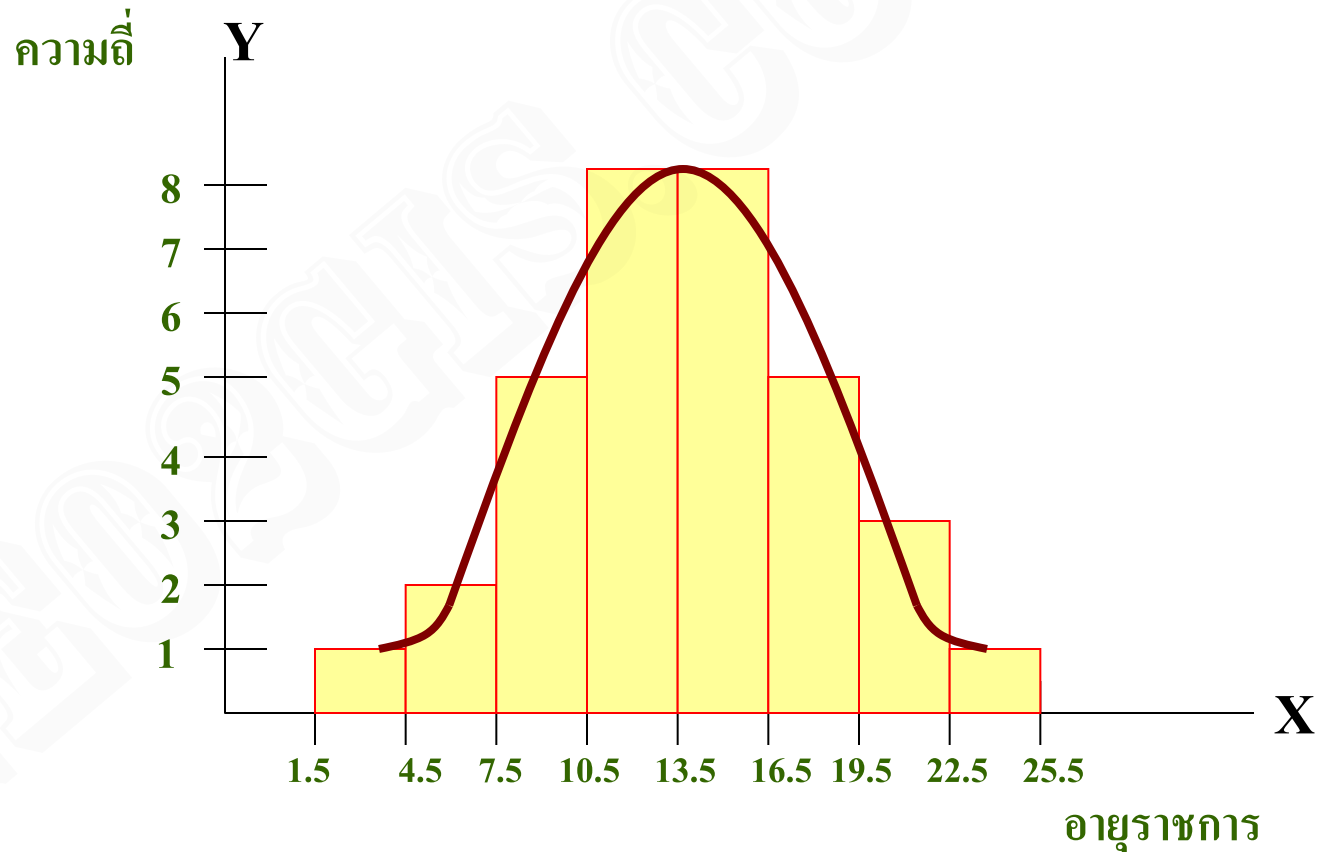
อายุราชการ	ความถี่
2 - 4	1
5 - 7	2
8 - 10	5
11 - 13	8
14 - 16	8
17 - 19	5
20 - 22	3
23 - 25	1



สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

การแจกแจงความถี่ด้วยแผนภูมิหรือกราฟ

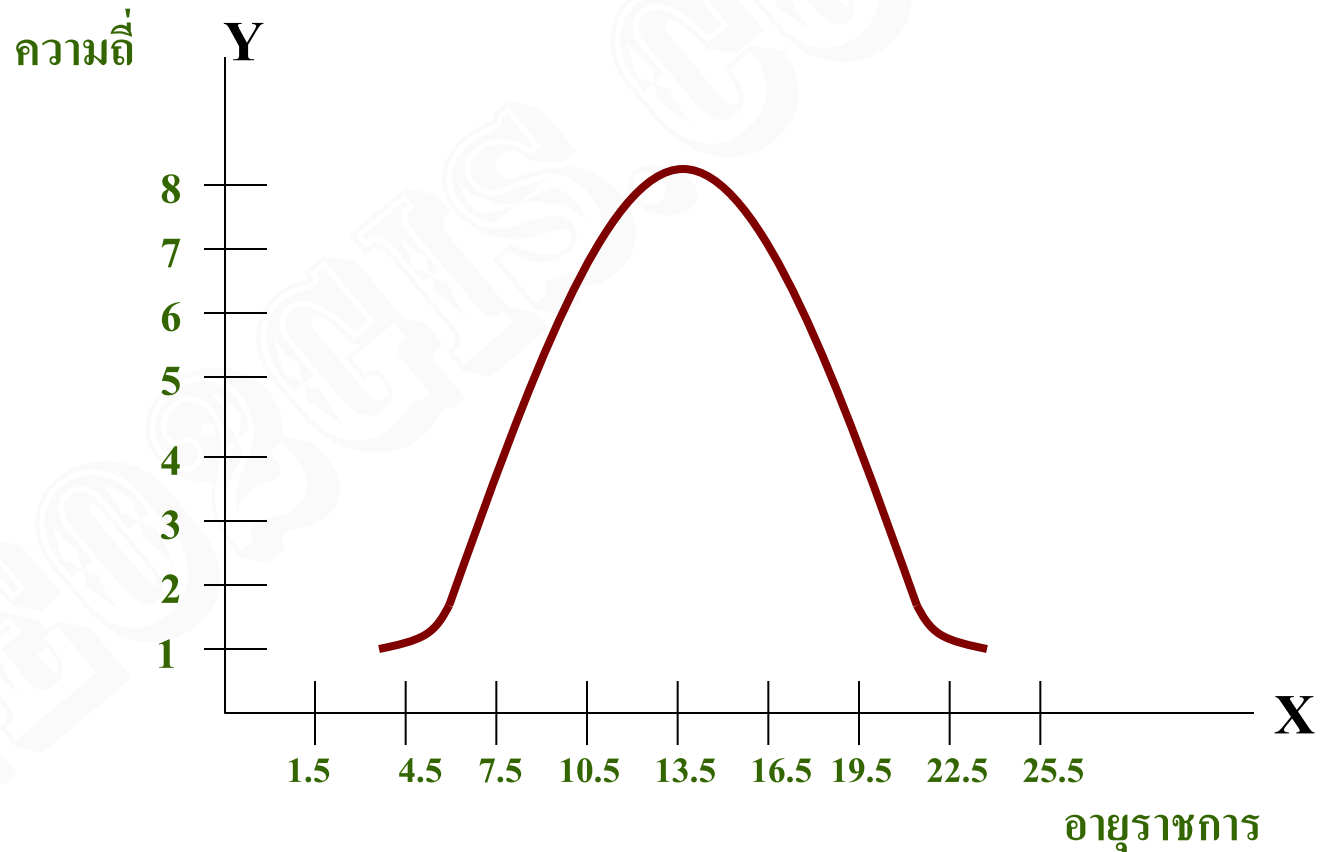
อายุราชการ	ความถี่
2 - 4	1
5 - 7	2
8 - 10	5
11 - 13	8
14 - 16	8
17 - 19	5
20 - 22	3
23 - 25	1



สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

การแจกแจงความถี่ด้วยแผนภูมิหรือกราฟ

อายุราชการ	ความถี่
2 - 4	1
5 - 7	2
8 - 10	5
11 - 13	8
14 - 16	8
17 - 19	5
20 - 22	3
23 - 25	1

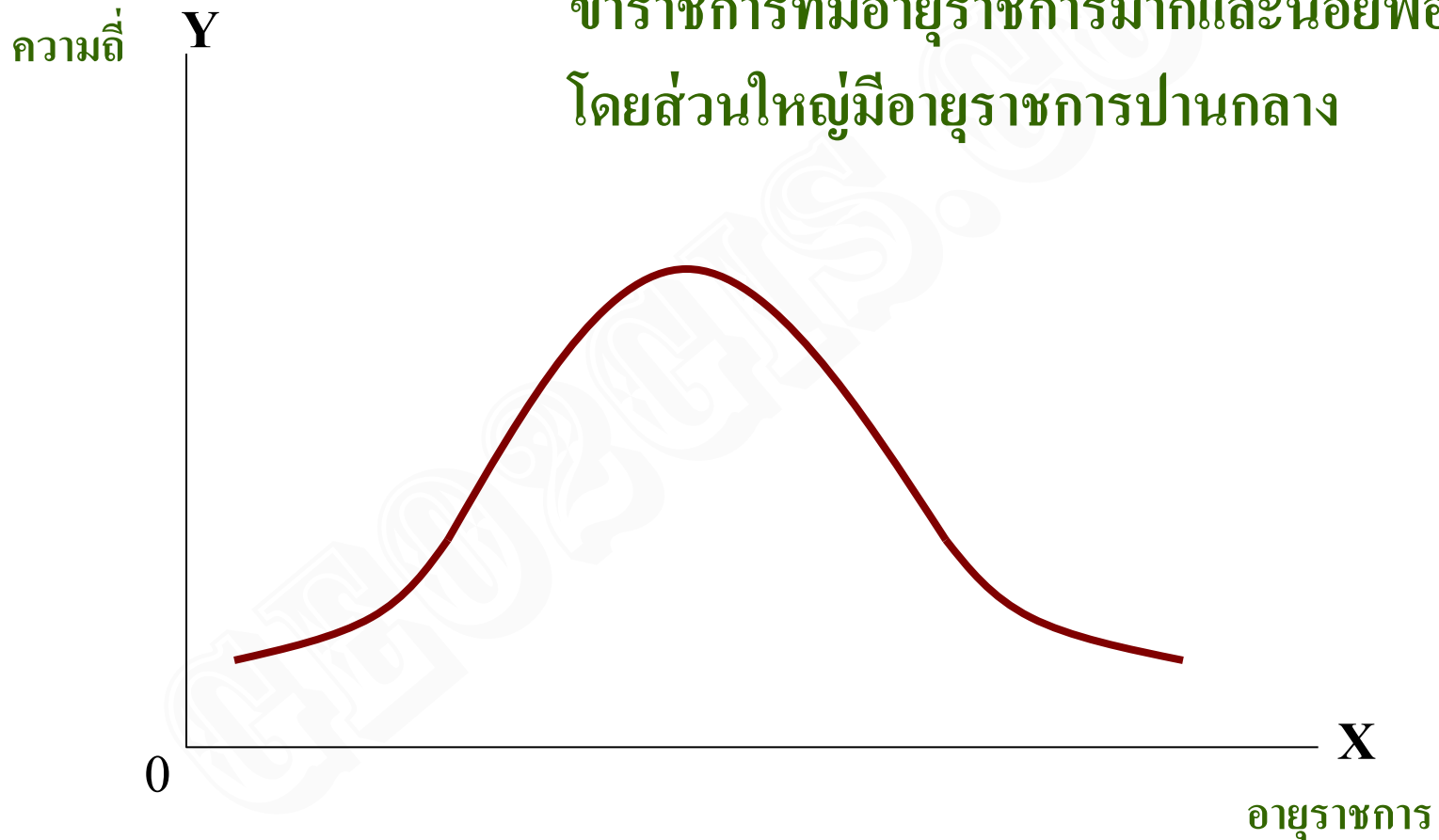


สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

โค้งปกติ (Normal Curve)

โค้งความถี่

ข้าราชการที่มีอายุราชการมากและน้อยพอ ๆ กัน
โดยส่วนใหญ่มีอายุราชการปานกลาง



สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

โค้งเบ้ทางบวก (Positively Skewed)



สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

โค้งเบ้ทางลบ (Negatively Skewed)



สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

โค้งความถี่

โค้งรูปตัวเจ (J-Shaped)

ความถี่

Y

ข้าราชการที่มีอายุราชการน้อยมีจำนวนน้อย
ข้าราชการที่มีอายุราชการมากมีจำนวนมาก

0

X

อายุราชการ



สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

โค้งความถี่

โค้งรูปตัวยู (U-Shaped)



สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

โค้งสองยอด (Bimodal Curve)



สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

โค้งเบ้ทางบวก (Positively Skewed)



สถิติพื้นฐานสำหรับการวิเคราะห์เชิงปริมาณ

โค้งเบ้ทางลบ (Negatively Skewed)

ความถี่

Y

ข้าราชการส่วนใหญ่ที่มีอายุราชการมาก
ข้าราชการที่มีอายุราชการน้อยมีจำนวนน้อย

0

X

อายุราชการ



ความเบ้ (Skewness)



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

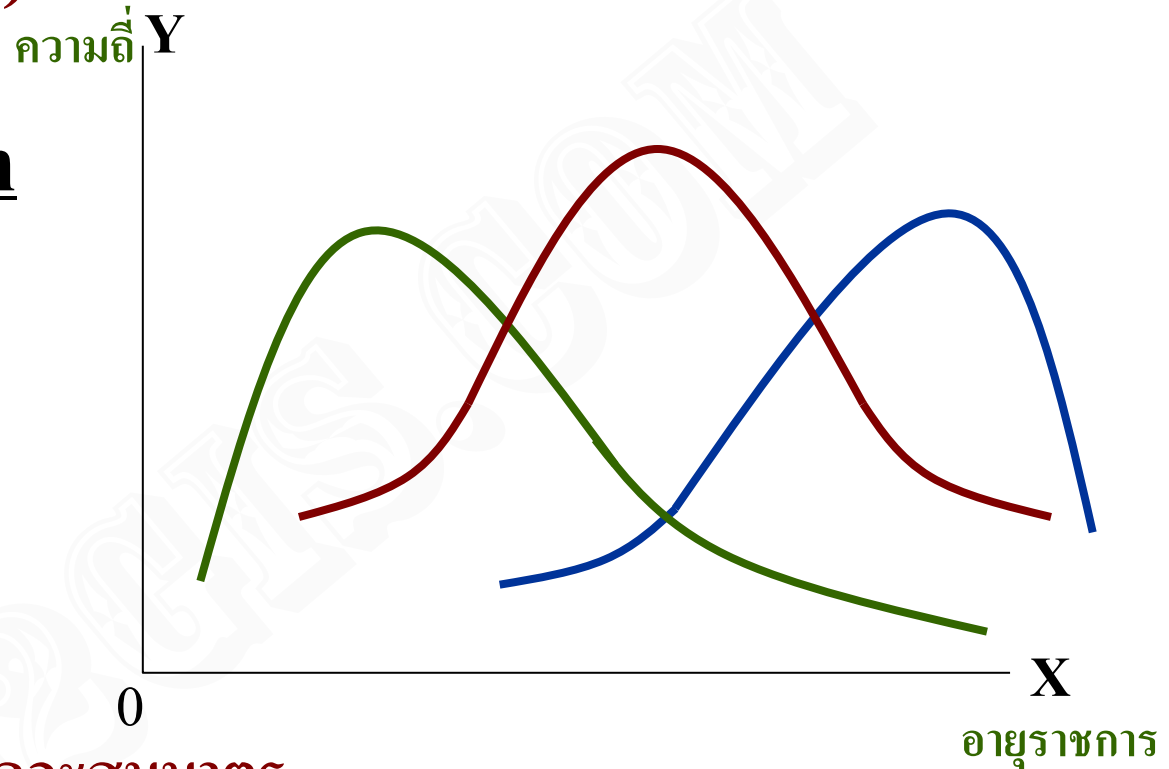
วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

ความเบ้ (Skewness)

$$Sk = \frac{\bar{X} - Mdn}{Mdn}$$



Sk : ไม่มีหน่วย

Sk = 0 : การกระจายข้อมูลจะสมมาตร

Sk มีค่า + : การกระจายข้อมูลเบ้ขวา

Sk มีค่า - : การกระจายข้อมูลเบ้ซ้าย



ความโด่ง (Kurtosis)



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

ความโด่ง (Kurtosis)

$$\text{Kur} = \frac{\left[\frac{\sum (x - \bar{x})^4}{N} \right]}{(SD^2)^2}$$

Kur = 3 : ข้อมูลแจกแจงแบบโค้งปกติ
ลักษณะโค้งมีค่าความโด่งปกติ

Kur < 3 : ข้อมูลมีความแปรปรวนมาก
ลักษณะโค้งจะแบนมีความโด่งน้อย

Kur > 3 : ข้อมูลมีความแปรปรวนน้อย
ลักษณะโค้งจะแบนมีความโด่งมาก

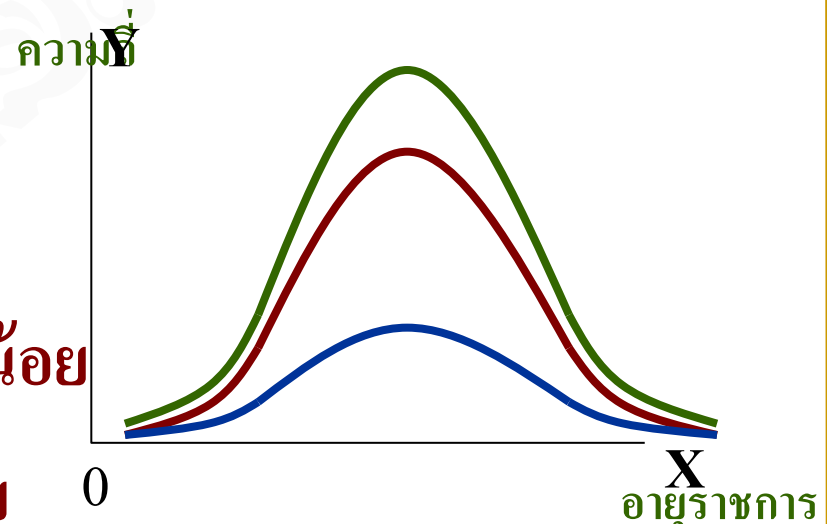
Kur : ค่าความโด่ง

X : ค่าคะแนน

$\bar{X}$: ค่าเฉลี่ย

N : ขนาดตัวอย่าง

SD² : ค่าความแปรปรวน



ตัวอย่างโจทย์



สถิติและฐานข้อมูลสำหรับนักบริหาร (Statistics and database for administrators)

วิทยาลัยการจัดการเพื่อการพัฒนา มหาวิทยาลัยทักษิณ

นายอนุสร พุ่มพวง

anu.pumpuang@gmail.com

1) เมื่อคะแนนรายงานของนิสิตแบ่งเป็น 3 ระดับ คือ สูง ปานกลาง ต่ำ ซึ่งมีคะแนน ดังนี้

11	12	11	13	14	15	10	8	12	5	15	10	17	11
18	11	7	11	6	11	17	7	19	7	12	13	15	5

จงสร้างตารางแจกแจงความถี่และคำนวณหาค่าเฉลี่ยเลขคณิต (Mean)

2) เมื่อคะแนนรายงานของนิสิตมีคะแนน ดังนี้

11 12 11 13 14 15 10 8 12 5
 15 10 17 11

จงคำนวณหาค่าเฉลี่ยเลขคณิต (Mean) ค่ามัธยฐาน (Median) และค่าฐานนิยม (Mode)

