

การวิเคราะห์ข้อมูลทางด้านสถิติ

โดยใช้โปรแกรม

Geogebra



ดำรงค์ ทิพย์โยธา

การวิเคราะห์ข้อมูลทางด้านสถิติ

โดยใช้โปรแกรม

Geogebra

รองศาสตราจารย์ ดำรงค์ ทิพย์โยธา

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ชื่อหนังสือ : การวิเคราะห์ข้อมูลทางด้านสถิติโดยใช้โปรแกรม Geogebra

ชื่อผู้แต่ง : รองศาสตราจารย์ ดำรงค์ ทิพย์โยธา

จัดทำโดย : รองศาสตราจารย์ ดำรงค์ ทิพย์โยธา

พิมพ์ครั้งที่ : 1

เดือนและปีที่ พ.ศ. ที่จัดพิมพ์ : สิงหาคม พ.ศ. 2563

เลขมาตรฐานสากลประจำหนังสืออิเล็กทรอนิกส์

ISBN (e-book) : 978-616-572-070-0

พิมพ์ที่ : ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์

จุฬาลงกรณ์มหาวิทยาลัย

คำนำ

หนังสือ การวิเคราะห์ข้อมูลทางด้านสถิติโดยใช้โปรแกรม Geogebra เป็นคู่มือในการใช้งานโปรแกรม Geogebra เพื่อวิเคราะห์ข้อมูลทางสถิติ เนื้อหาภายในเล่มประกอบด้วย การสร้างแฟ้มข้อมูล การคำนวณค่าสถิติเบื้องต้น เช่น ค่าเฉลี่ย ค่ามัธยฐาน ค่าความแปรปรวน ส่วนเบี่ยงเบนมาตรฐาน ฯลฯ การนำเสนอข้อมูลในรูปแบบการแจกแจงความถี่ การนำเสนอในรูปแบบกราฟ เขียนแผนภาพการกระจายของข้อมูล วิเคราะห์ข้อมูลทางด้านสถิติ ทำตาราง ANOVA และหาค่าความน่าจะเป็นของตัวแปรสุ่ม

โปรแกรม Geogebra ช่วยคำนวณค่าสถิติและวิเคราะห์ข้อมูลทางสถิติได้เช่น หาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าพารามิเตอร์ ทดสอบสมมติฐาน $H_0 : \mu = \mu_0$, $H_0 : \mu_1 - \mu_2 = d_0$, $H_0 : \mu_D = d_0$, $H_0 : \sigma^2 = \sigma_0^2$, $H_0 : \frac{\sigma_1^2}{\sigma_2^2} = k$ ทดสอบภาวะสารูปสนิทธิ ทดสอบสมมติฐานว่าข้อมูลเป็นอิสระกันหรือไม่ หาสมการเส้นถดถอยเชิงเส้นตรงเชิงเดียว (Simple Linear Regression) สหสัมพันธ์ (Correlation) หาสัมประสิทธิ์สหสัมพันธ์และสัมประสิทธิ์การกำหนด หารูปแบบความสัมพันธ์ที่ดีที่สุดสำหรับข้อมูล และวิเคราะห์ความแปรปรวน ในการวิเคราะห์ข้อมูลบางกรณีจำเป็นต้องใช้โปรแกรม Excel Mathcad Mathematica มาช่วยในการคำนวณเพื่อช่วยเสริมให้เนื้อหา มีความสมบูรณ์มากขึ้น

เนื้อหาภายในเล่มจะแสดงขั้นตอนการทำงานทางสถิติ และมีเหตุผลของทฤษฎีความน่าจะเป็นและสถิติประกอบการทำงาน และแสดงสูตรสถิติซึ่งเป็นที่มาของค่าสถิติต่าง ๆ ที่ Geogebra คำนวณมาให้ สรุปผลการวิเคราะห์ข้อมูลและแปลความหมายทางสถิติเพื่อนำไปใช้งาน เพื่อความสะดวกและรวดเร็วในการศึกษาผู้อ่านทุกท่านสามารถ Download ข้อมูลตัวอย่าง โดยเข้าไปที่

http://pioneer.netserv.chula.ac.th/~tdumrong/2301286/datasps/stat_geogebra.rar

ผู้เขียนหวังว่าหนังสือเล่มนี้จะช่วยให้ผู้อ่านทุกท่านสามารถนำโปรแกรม Geogebra ไปใช้วิเคราะห์ข้อมูลได้เป็นอย่างดี และขอขอบคุณผู้อ่านทุกท่านที่ได้ติดตามผลงานของผู้เขียนมาโดยตลอด

รองศาสตราจารย์ ดำรงค์ ทิพย์โยธา

สิงหาคม 2563

สารบัญ		หน้า
บทที่ 1	การวิเคราะห์ข้อมูลทางสถิติโดยใช้โปรแกรม Geogebra	1 – 44
	1.1 ตัวอย่างการวิเคราะห์ข้อมูลทางสถิติโดยใช้โปรแกรม Geogebra	1
	1.2 การสร้างแฟ้มข้อมูล	13
	1.3 การหาค่าสถิติเบื้องต้น	20
บทที่ 2	การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าพารามิเตอร์	45 – 200
	2.1 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าเฉลี่ย μ	45
	2.2 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่าง ของค่าเฉลี่ย $\mu_1 - \mu_2$	100
	2.3 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าเฉลี่ยของผลต่าง μ_D กรณีประชากร 2 ชุดไม่เป็นอิสระต่อกัน (ข้อมูลเป็นคู่)	165
	2.4 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของความแปรปรวน σ^2	190
	2.5 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ $\frac{\sigma_1^2}{\sigma_2^2}$	195
บทที่ 3	การทดสอบสมมติฐาน	201 – 424
	3.1 การทดสอบสมมติฐาน $H_0 : \mu = \mu_0$	202
	3.2 การทดสอบสมมติฐาน $H_0 : \mu_1 - \mu_2 = d_0$ กรณีประชากร 2 ชุดเป็นอิสระต่อกัน	256
	3.3 การทดสอบสมมติฐาน $H_0 : \mu_D = d_0$ กรณีประชากร 2 ชุดไม่เป็นอิสระต่อกัน (ข้อมูลเป็นคู่)	340
	3.4 การทดสอบสมมติฐาน $H_0 : \sigma^2 = \sigma_0^2$	371
	3.5 การทดสอบสมมติฐาน $\frac{\sigma_1^2}{\sigma_2^2} = k$	378
	3.6 การทดสอบภาวะสารถูปสนิทธิ (Test of Goodness Of Fit)	391
	3.7 การทดสอบสมมติฐานว่าข้อมูลเป็นอิสระต่อกันหรือไม่	407

บทที่ 4	สหสัมพันธ์และการถดถอย	425 – 444
4.1	การหาสมการเส้นถดถอยเชิงเส้นตรงเชิงเดียว (Simple Linear Regression) และสหสัมพันธ์ (Correlation)	427
บทที่ 5	การวิเคราะห์ความแปรปรวน	445 – 458
5.1	การวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียว (One – Way ANOVA, Simple – Factor ANOVA)	113
บทที่ 6	การเปิดตารางสถิติหาค่าความน่าจะเป็น	459 – 496
6.1	การแจกแจงปกติ (Normal Distribution) และการแจกแจงปกติมาตรฐาน (Standard Normal Distribution)	460
6.2	การแจกแจงที (Student's t Distribution)	465
6.3	การแจกแจงไคสแควร์ (χ^2 : Chi – Squared Distribution)	470
6.4	การแจกแจงเอฟ (F – Distribution)	475
6.5	การแจกแจงทวินาม (Binomial Distribution)	480
6.6	การแจกแจงไฮเพอร์จีโอเมตริก (Hypergeometric Distribution)	485
6.7	การแจกแจงปัวส์ซอง (Poisson Distribution)	491

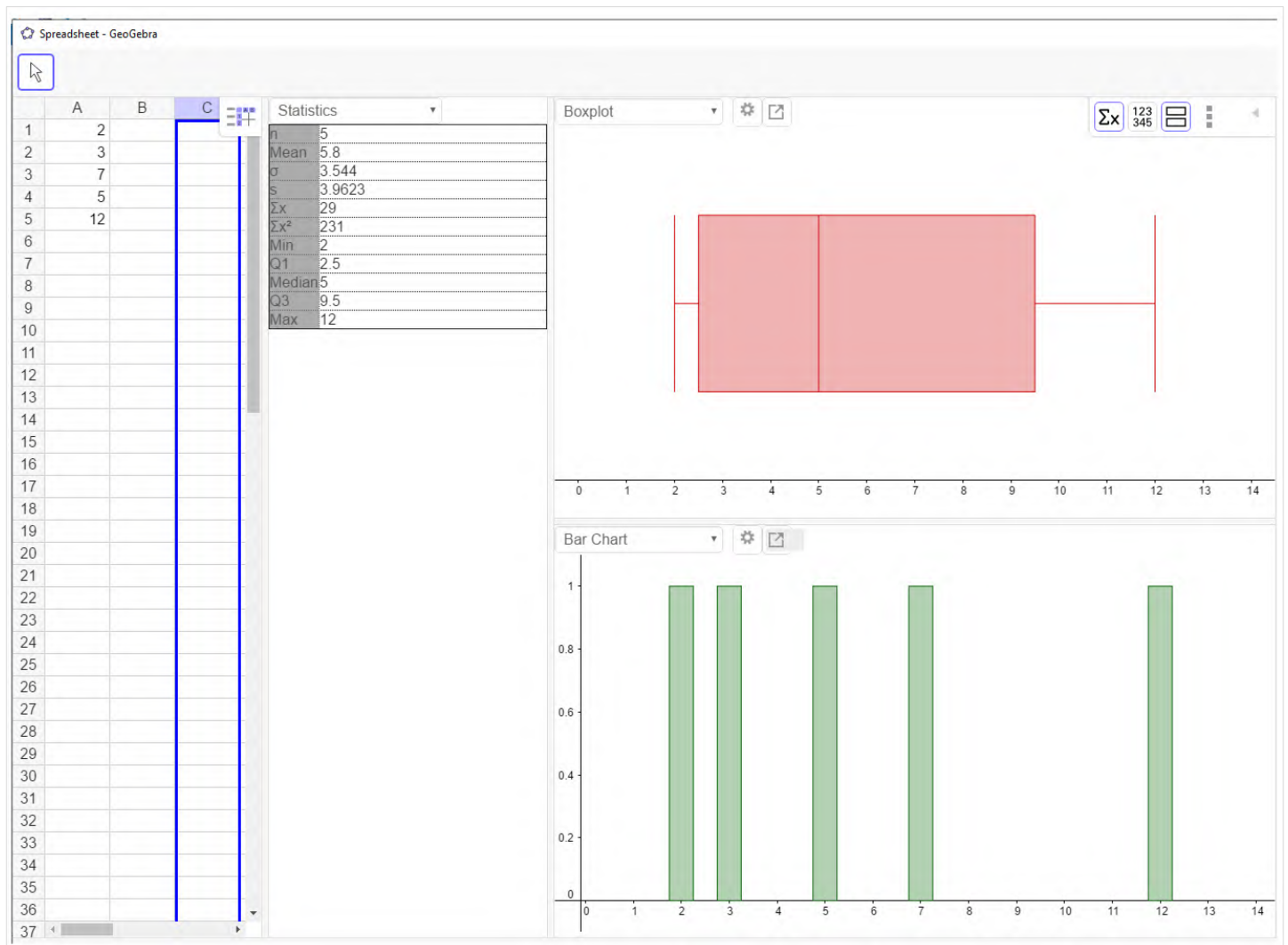
บทที่ 1

การวิเคราะห์ข้อมูลทางสถิติโดยใช้โปรแกรม Geogebra

โปรแกรม Geogebra มีความสามารถในการทำงานทางด้านสถิติเช่น การนำเสนอข้อมูล หาค่าสถิติเบื้องต้น และวิเคราะห์ข้อมูลทางด้านสถิติได้

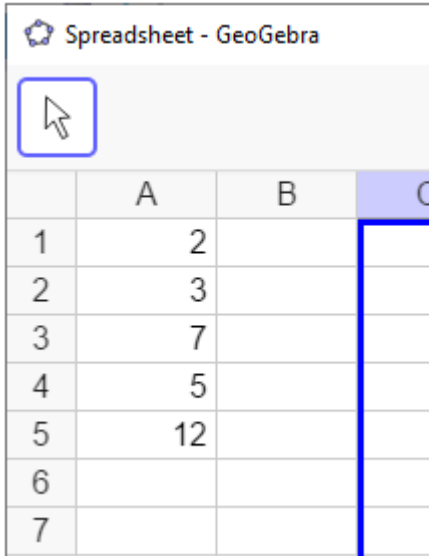
1.1 ตัวอย่างการวิเคราะห์ข้อมูลทางสถิติโดยใช้โปรแกรม Geogebra

ตัวอย่างการหาค่าสถิติเบื้องต้น เช่น ค่าเฉลี่ย มัธยฐาน ฐานนิยม ส่วนเบี่ยงเบนมาตรฐาน



ข้อมูลจำนวน 5 ตัวคือ 2, 3, 5, 7 และ 12

Spreadsheet - GeoGebra



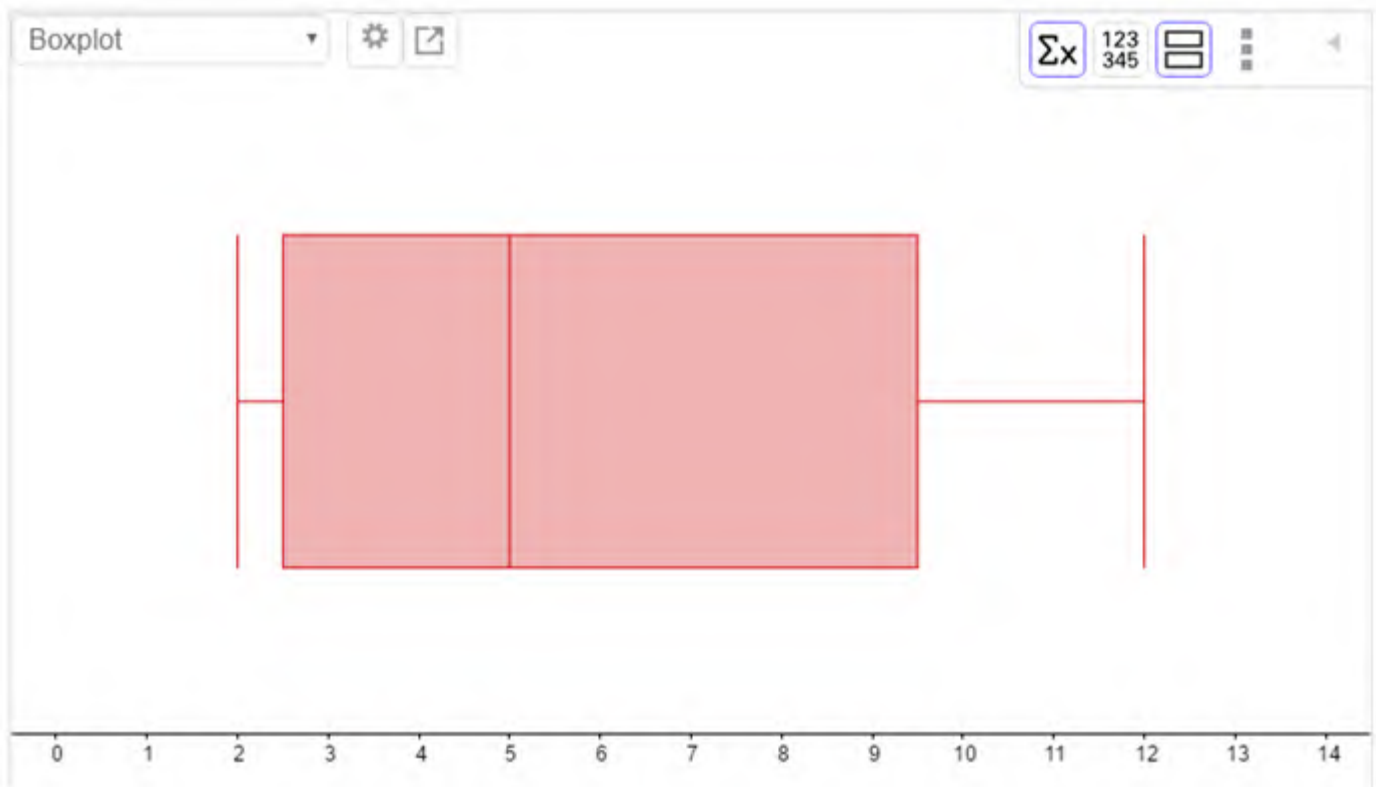
	A	B	C
1	2		
2	3		
3	7		
4	5		
5	12		
6			
7			

คำนวณค่าสถิติเบื้องต้น

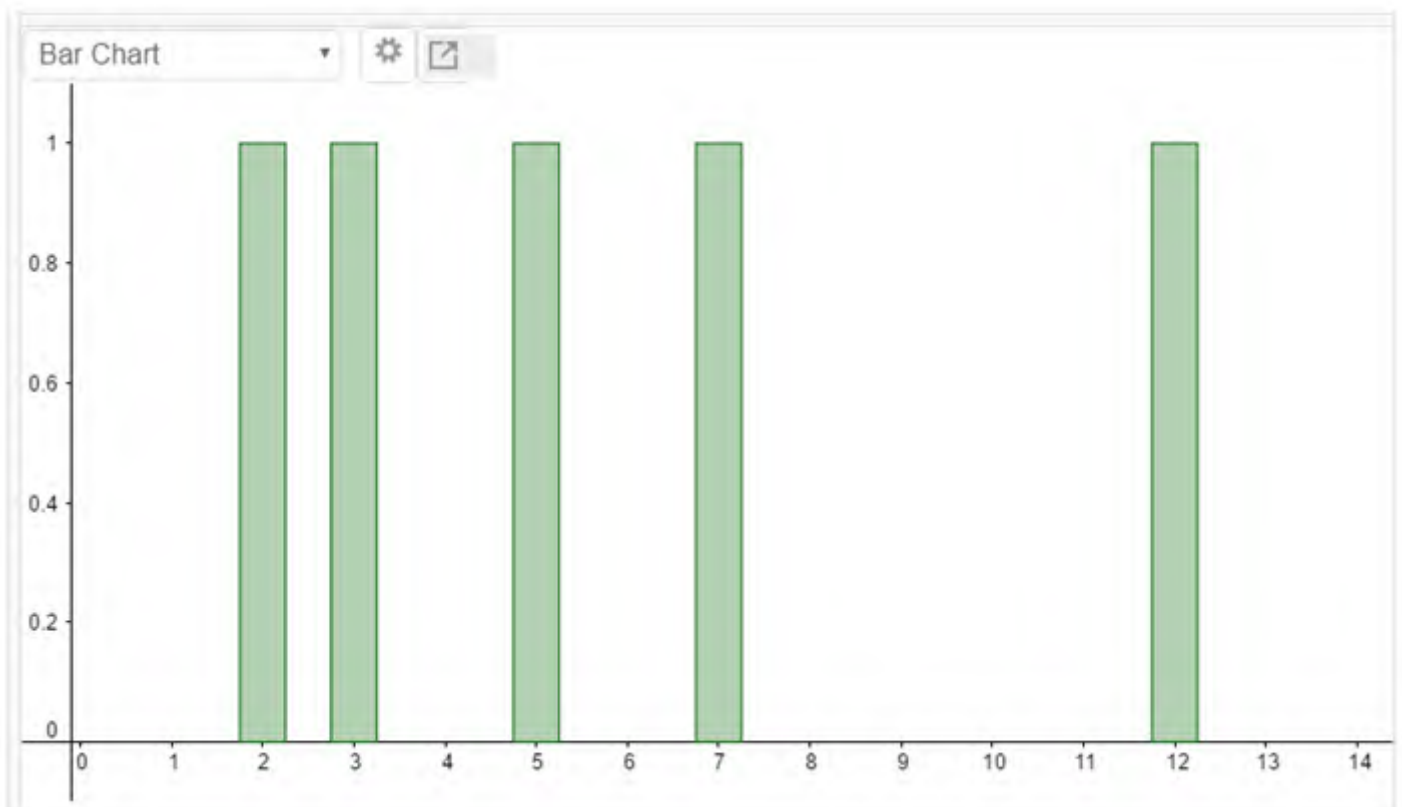
Statistics ▼

n	5
Mean	5.8
σ	3.544
s	3.9623
Σx	29
Σx^2	231
Min	2
Q1	2.5
Median	5
Q3	9.5
Max	12

แผนภาพ Box-Plot



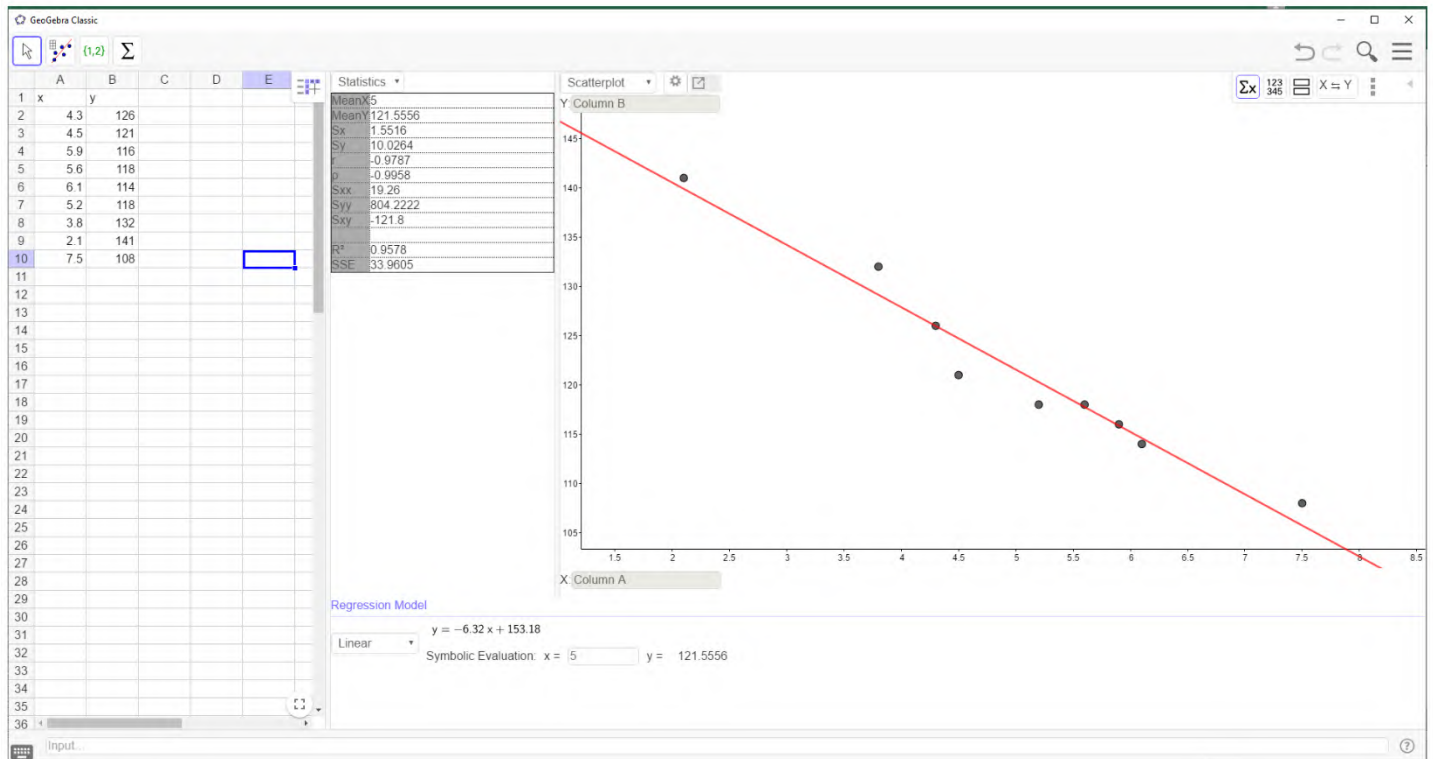
แผนภาพ Bar-Graph



การหาความสัมพันธ์เชิงเส้นตรงของข้อมูล ตัวอย่างเช่นการหาความสัมพันธ์เชิงเส้นตรง $\mu_{Y|x} = \alpha + \beta \times$ ระหว่าง ปริมาณน้ำฝนที่ตก ($x = \text{rain}$ หน่วย : มิลลิเมตร) กับ ปริมาณอากาศเป็นพิษที่ถูกกำจัดออกไป ($y = \text{air}$ หน่วย : ไมโครกรัมต่อลูกบาศก์เมตร) จากข้อมูลต่อไปนี้

rain	4.3	4.5	5.9	5.6	6.1	5.2	3.8	2.1	7.5
air	126	121	116	118	114	118	132	141	108

ผลการวิเคราะห์ข้อมูล



การนำเข้าข้อมูล

GeoGebra Classic





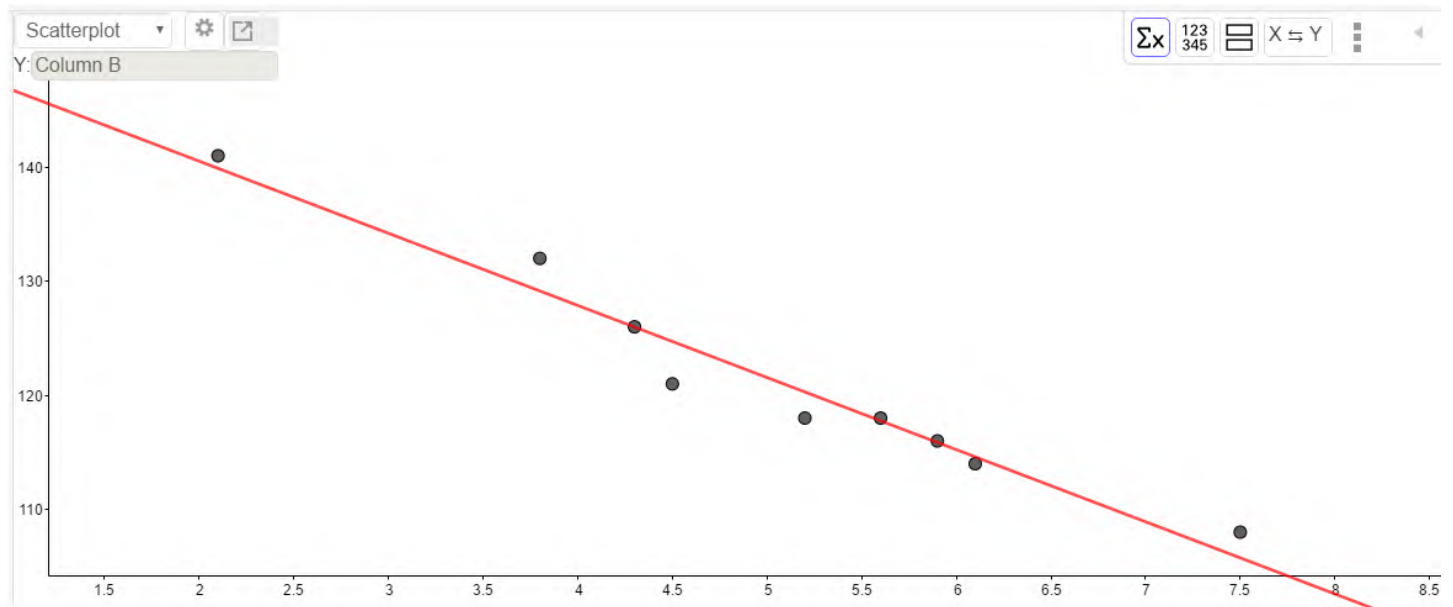

	A	B	C	D
1	rain	air		
2	4.3	126		
3	4.5	121		
4	5.9	116		
5	5.6	118		
6	6.1	114		
7	5.2	118		
8	3.8	132		
9	2.1	141		
10	7.5	108		

ผลการคำนวณเบื้องต้น

Statistics ▾

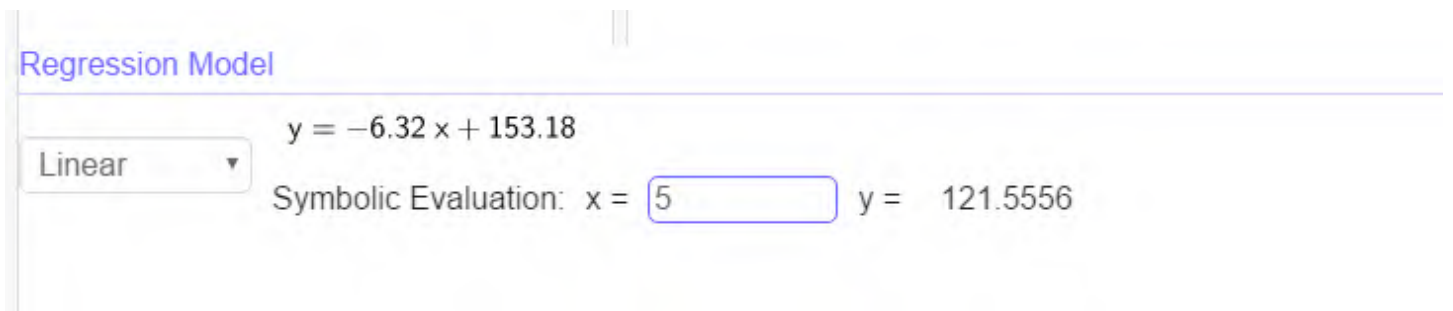
MeanX	5
MeanY	121.5556
Sx	1.5516
Sy	10.0264
r	-0.9787
p	-0.9958
Sxx	19.26
Syy	804.2222
Sxy	-121.8
R ²	0.9578
SSE	33.9605

แผนภาพการกระจายของข้อมูลพร้อมสมการเส้นถดถอย



ตัวแปรอิสระ x (rain) ตัวแปรตาม y (air)

สมการเส้นถดถอย $y = -6.32x + 153.18$ พร้อมการพยากรณ์ค่า y เมื่อ $x = 5$



ทดสอบความเป็นอิสระตัวอย่างเช่น ในการศึกษาความสัมพันธ์ระหว่างการนับถือศาสนาและถิ่นที่อยู่อาศัย เลือกคนมา 2 กลุ่มโดยวิธีสุ่ม กลุ่มหนึ่งอาศัยอยู่แถบชายฝั่งทะเลตะวันออกของประเทศหนึ่ง อีกกลุ่มหนึ่งอาศัยอยู่แถบชายฝั่งทะเลตะวันตก ปรากฏจำนวนผู้นับถือศาสนาต่าง ๆ มีดังนี้

	โปรเตสตัน	คาทอลิก	ยิว	รวม
ฝั่งตะวันออก	182	215	203	600
ฝั่งตะวันตก	154	136	110	400
รวม	336	351	313	1000

จงทดสอบว่าการนับถือศาสนาและถิ่นที่อยู่อาศัยมีความสัมพันธ์กันหรือไม่ กำหนดระดับนัยสำคัญ 0.05

วิธีทำ ให้ เหตุการณ์ A เป็นเหตุการณ์การนับถือศาสนา

ให้ เหตุการณ์ B เป็นเหตุการณ์ถิ่นที่อยู่อาศัย

การทดสอบสมมติฐานว่าการนับถือศาสนาและถิ่นที่อยู่อาศัยมีความสัมพันธ์กันหรือไม่ คือการทดสอบว่าเหตุการณ์ A และ B เป็นอิสระต่อกัน

ขั้นที่ 1 กำหนดสมมติฐานหลัก H_0 : เหตุการณ์ A และ B เป็นอิสระต่อกัน

กำหนดสมมติฐานทางเลือก H_1 : เหตุการณ์ A และ B ไม่เป็นอิสระต่อกัน

ขั้นที่ 2 กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3 สุ่มตัวอย่าง

ขั้นที่ 4 เลือกใช้ค่าสถิติไคสแควร์

ขั้นที่ 5 คำนวณค่าสถิติจากตัวอย่าง $\chi^2_{\text{คำนวณ}}$

	โปรเตสตัน	คาทอลิก	ยิว	รวม
ฝั่งตะวันออก	182	215	203	600
ฝั่งตะวันตก	154	136	110	400
รวม	336	351	313	1000

ถ้า H_0 เป็นจริง ความถี่ที่คาดหวังได้คือ

$$e_{11} = \frac{R_1 C_1}{N} = \frac{(600)(336)}{1000} = 201.6$$

$$e_{12} = \frac{R_1 C_2}{N} = \frac{(600)(351)}{1000} = 134.4$$

$$e_{13} = \frac{R_1 C_3}{N} = \frac{(600)(313)}{1000} = 187.8$$

$$e_{21} = \frac{R_2 C_1}{N} = \frac{(400)(336)}{1000} = 134.4$$

$$e_{22} = \frac{R_2 C_2}{N} = \frac{(400)(351)}{1000} = 140.4$$

$$e_{23} = \frac{R_2 C_3}{N} = \frac{(400)(313)}{1000} = 125.2$$

$$\begin{aligned} \chi^2_{\text{คำนวณ}} &= \sum_{i=1}^2 \sum_{j=1}^3 \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \\ &= \frac{(182-201.6)^2}{201.6} + \frac{(215-210.6)^2}{210.6} + \frac{(203-187.8)^2}{187.8} \\ &\quad + \frac{(154-134.4)^2}{134.4} + \frac{(136-140.4)^2}{140.4} + \frac{(110-125.2)^2}{125.2} \\ &= 1.90555556 + 0.09192783 + 1.23024494 + 2.85833333 + 0.13789174 + 1.84536741 \\ &= 8.06932081 \end{aligned}$$

ขั้นที่ 6 การหาค่าวิกฤตและบริเวณวิกฤต

ระดับขั้นความเสรี $V = (r - 1)(c - 1) = (2 - 1)(3 - 1) = 2$

ค่าวิกฤตคือ $\chi^2_{0.05, 2} = 5.991$ บริเวณวิกฤตคือ $\chi^2 > 5.991$

Sig. (1-tailed) = $P(\chi^2 > 8.06932081) = 0.01769169$

ขั้นที่ 7 สรุปผล

แบบที่ 1 เพราะว่า $\chi^2_{\text{คำนวณ}}$ อยู่ในบริเวณวิกฤต เพราะฉะนั้น ปฏิเสธ H_0

แบบที่ 2 เพราะว่า Sig. (1-tailed) ไม่น้อยกว่า 0.05 เพราะฉะนั้น ปฏิเสธ H_0

เพราะฉะนั้นการนับถือศาสนาและถิ่นที่อยู่อาศัยมีความสัมพันธ์กัน ที่ระดับนัยสำคัญ 0.05



การคำนวณโดยใช้โปรแกรม Geogebra

chisq_test_independent_15May2020_1417 - GeoGebra

Distribution **Statistics**

ChiSquared Test

Rows 2
Columns 3

☒ Row % ☒ Column % ☒ Expected Count ☒ X² Contribution

	C1	C2	C3
	182	215	203
	201.6	210.6	187.8
	1.9056	0.0919	1.2302
	30.3333%	35.8333%	33.8333%
	54.1667%	61.2536%	64.8562%
R1			
	154	136	110
	134.4	140.4	125.2
	2.8583	0.1379	1.8454
	38.5%	34%	27.5%
	45.8333%	38.7464%	35.1438%
R2			
	336	351	313
	33.6%	35.1%	31.3%

Result

ChiSquared Test

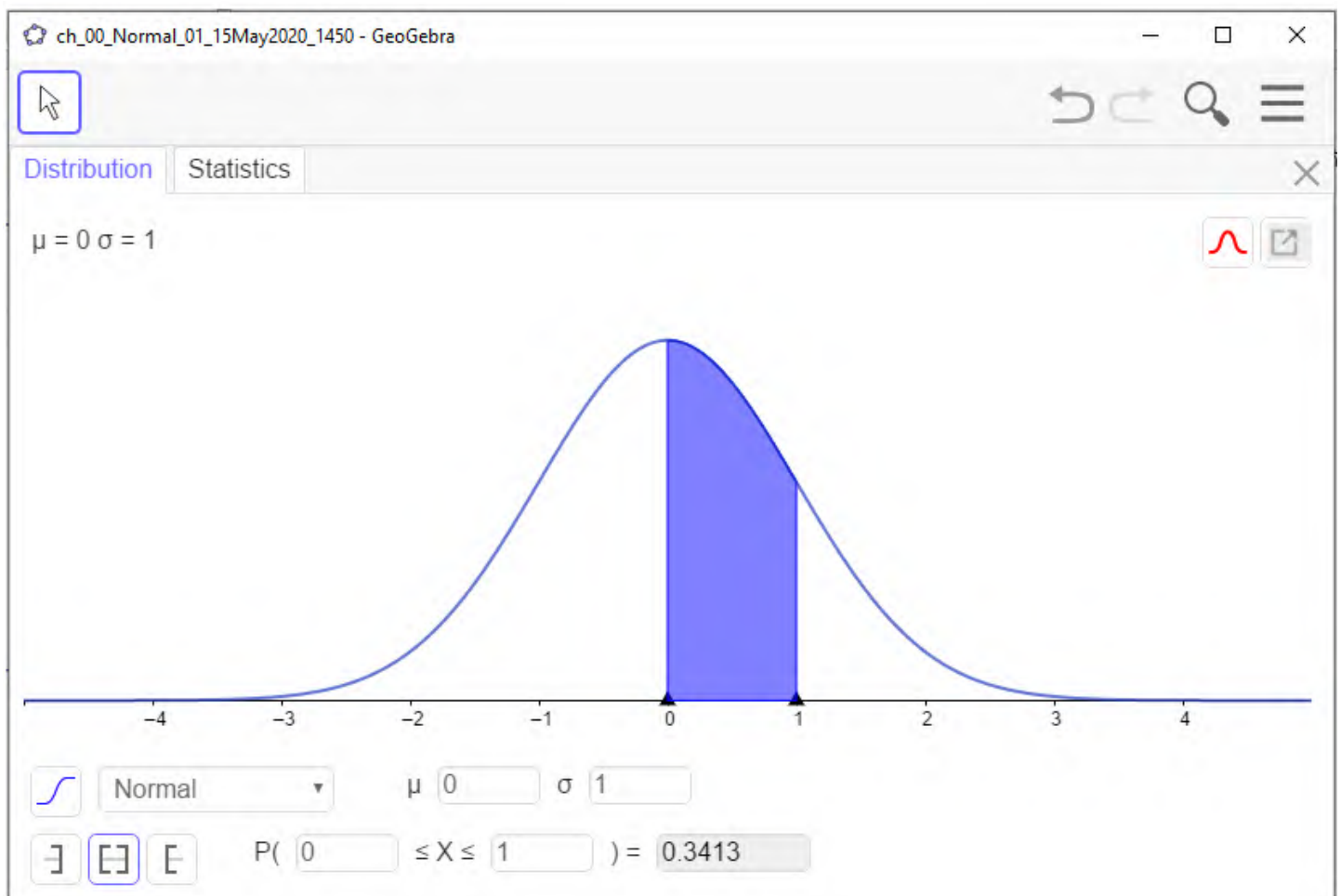
df 2
X² 8.0693
P 0.0177

$$\begin{aligned}
 \chi^2_{\text{คำนวณ}} &= \sum_{i=1}^2 \sum_{j=1}^3 \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \\
 &= \frac{(182 - 201.6)^2}{201.6} + \frac{(215 - 210.6)^2}{210.6} + \frac{(203 - 187.8)^2}{187.8} + \frac{(154 - 134.4)^2}{134.4} + \frac{(136 - 140.4)^2}{140.4} + \frac{(110 - 125.2)^2}{125.2} \\
 &= 1.9056 + 0.0919 + 1.2302 + 2.8583 + 0.1379 + 1.8454 \\
 &= 8.0693
 \end{aligned}$$

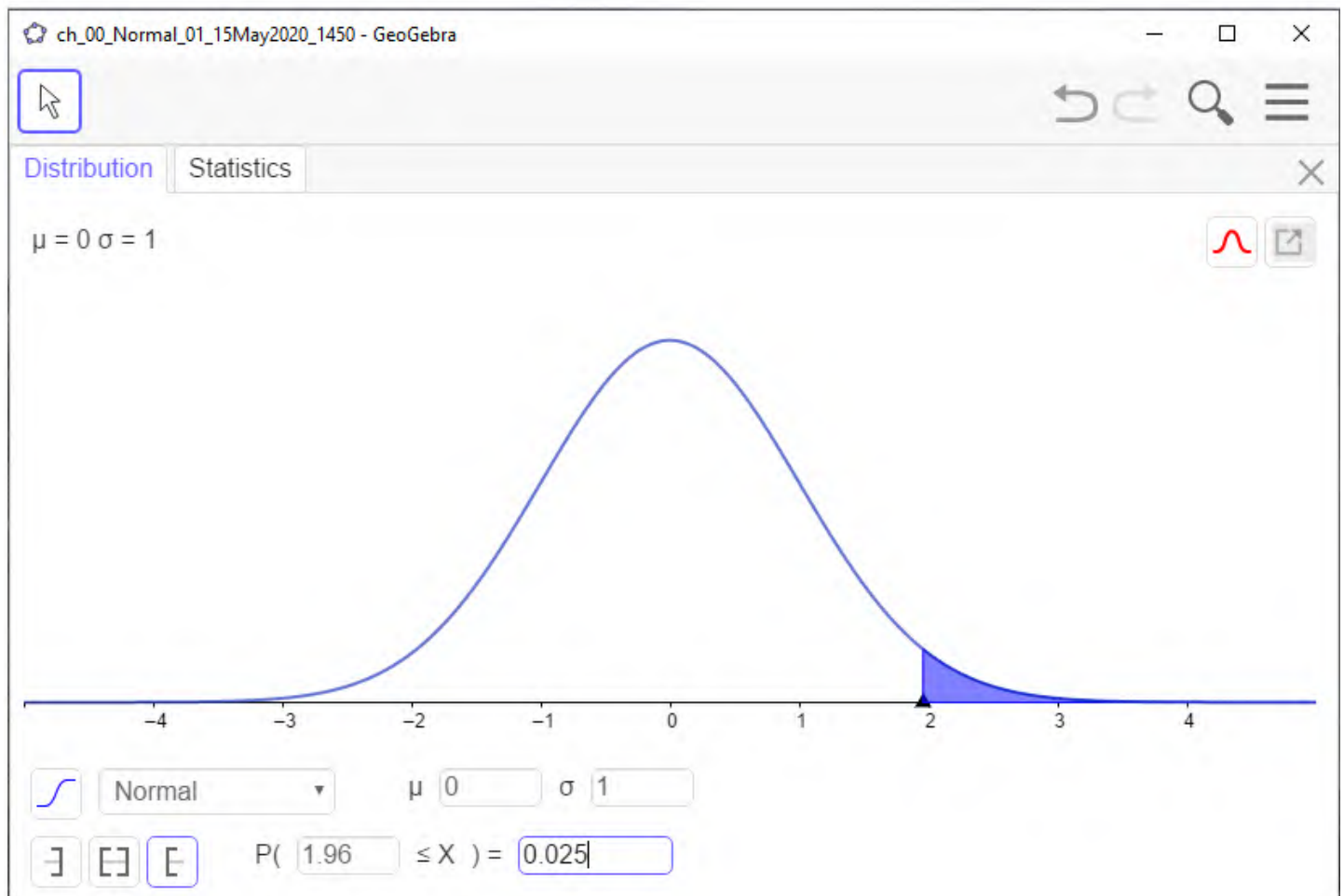
Sig. (1-tailed) = 0.0177

การแจกแจงความน่าจะเป็นปกติมาตรฐาน

กราฟของการแจกแจงความน่าจะเป็นของตัวแปรสุ่มปกติมาตรฐาน



$$P(0 < Z < 1) = 0.3413$$

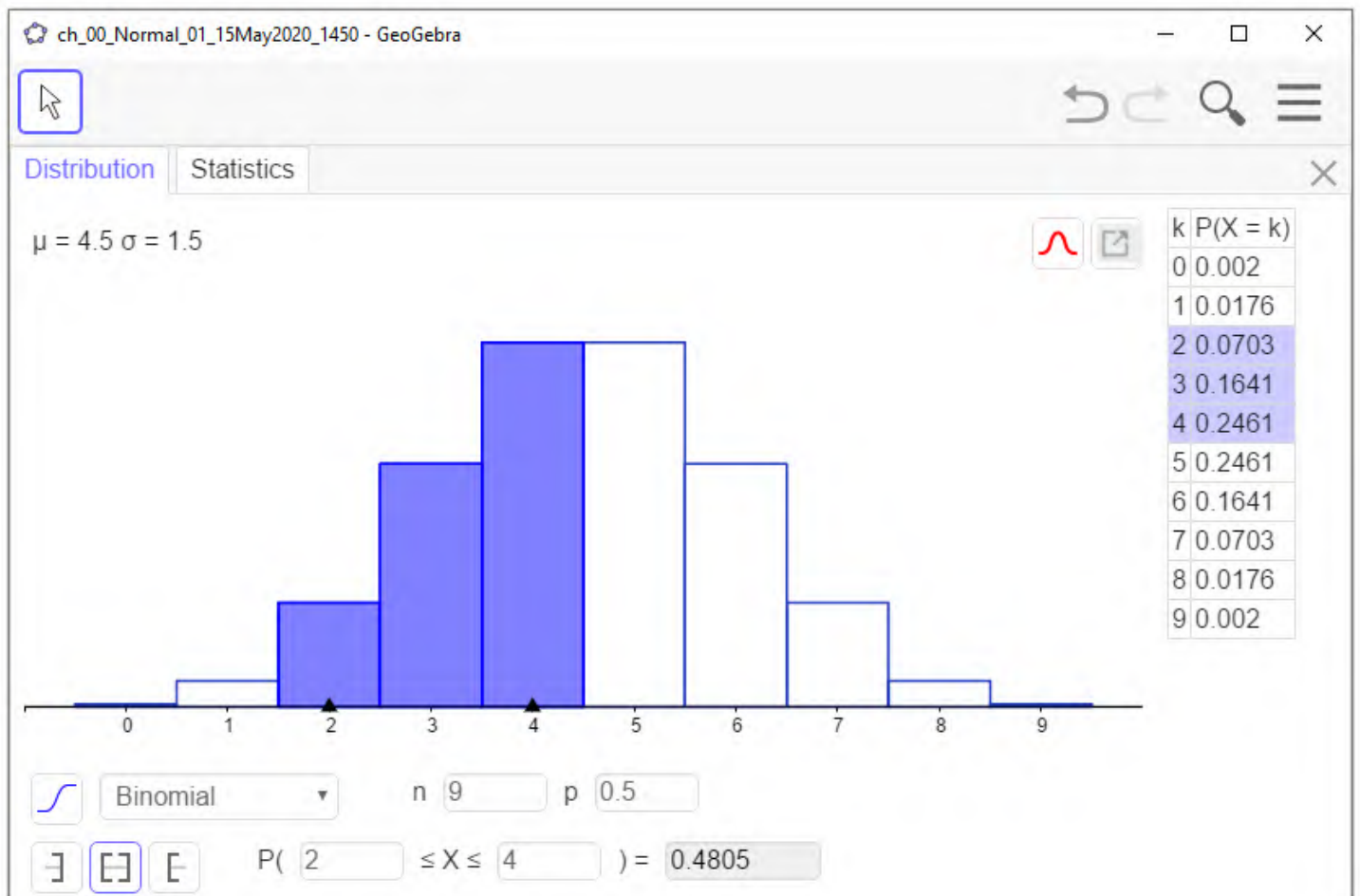


$$P(Z > 1.96) = 0.025$$

$$z_{0.025} = 1.96$$

การแจกแจงของตัวแปรสุ่มทวินาม

กราฟของการแจกแจงความน่าจะเป็นของตัวแปรสุ่มทวินาม



การแจกแจงความน่าจะเป็นทวินาม $n = 9$, $p = 0.5$, $P(2 \leq X \leq 4) = 0.4805$

1.2 การสร้างแฟ้มข้อมูล

สิ่งที่สำคัญของผู้ที่ต้องการวิเคราะห์ข้อมูลทางสถิติคือ การวางแผนเก็บข้อมูล การสร้างแบบสอบถาม การแปลความหมายแบบสอบถามเพื่อเป็นข้อมูลของการวิเคราะห์ด้วยโปรแกรม Geogebra ตัวอย่างเช่น บริษัทแห่งหนึ่งต้องการเก็บรวบรวมข้อมูลและวิเคราะห์ข้อมูลของพนักงานเกี่ยวกับ เพศ อายุ ระดับการศึกษา สถานภาพ เงินเดือน ระดับคะแนนความสามารถ และ เงินตอบแทนประจำปี จึงทำการสอบถามข้อมูลด้วยแบบสอบถามดังนี้

แบบสอบถามข้อมูลพนักงาน

สำหรับเจ้าหน้าที่กรอกข้อมูล

1. เลขประจำตัว.....
2. เพศ ☐ ชาย ☐ หญิง ☐
3. อายุ ปี
4. ระดับการศึกษา ☐ ต่ำกว่าระดับปริญญาตรี ☐ จบระดับปริญญาตรี ☐
☐ จบระดับปริญญาโท ☐ จบระดับปริญญาเอก
5. สถานภาพ ☐ โสด ☐ แต่งงานแล้ว ☐
☐ เป็นหม้าย ☐ หย่าร้าง
6. เงินเดือนบาท
7. ระดับคะแนนความสามารถ.....
8. เงินตอบแทนประจำปี.....บาท

ข้อกำหนดในการสร้างแฟ้มข้อมูล

จากแบบสอบถามที่ผู้ที่ต้องการวิเคราะห์ข้อมูลเก็บรวบรวมมา เมื่อต้องการจะทำการเป็นข้อมูลสำหรับ Geogebra ต้องทำการกำหนดค่าต่าง ๆ เช่น ชื่อแฟ้มข้อมูล (file name) ชื่อตัวแปร (variable name) กำหนดค่าข้อมูลที่ไม่สมบูรณ์ (missing value) คำอธิบายความหมายของค่าตัวแปร (value label)

ข้อกำหนดของแฟ้มข้อมูลที่เรากำลังต้องการเป็นดังนี้

1. เลขประจำตัว ชื่อตัวแปร id
กำหนดชนิดของข้อมูล จำนวนเต็ม
ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์ ไม่มี
คำอธิบายความหมายของค่าตัวแปร ไม่มี
2. เพศ ชื่อตัวแปร sex
กำหนดชนิดของข้อมูล จำนวนเต็ม
ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์ 9
คำอธิบายความหมายของค่าตัวแปร 1 Male 2 Female
3. อายุ ชื่อตัวแปร age
กำหนดชนิดของข้อมูล จำนวนเต็ม
ข้อมูลที่ไม่สมบูรณ์ให้เว้นว่าง
4. ระดับการศึกษา ชื่อตัวแปร educ
กำหนดชนิดของข้อมูล จำนวนเต็ม
ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์ 9
คำอธิบายความหมายของค่าตัวแปร 1 Under_gr 2 Graduate 3 Post_gr 4 Doctor
5. สถานภาพ ชื่อตัวแปร status
กำหนดชนิดของข้อมูล จำนวนเต็ม
ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์ 9
คำอธิบายความหมายของค่าตัวแปร 1 Single 2 Married 3 Widow 4 Divorced
6. เงินเดือน ชื่อตัวแปร income
กำหนดชนิดของข้อมูล จำนวนจริงทศนิยม 2 ตำแหน่ง
ข้อมูลที่ไม่สมบูรณ์ให้เว้นว่าง

7. ระดับคะแนน ชื่อตัวแปร

grade

กำหนดชนิดของข้อมูล

จำนวนจริงทศนิยม 2 ตำแหน่ง

ข้อมูลที่ไม่สมบูรณ์ให้เว้นว่าง

8. เงินตอบแทนประจำปี ชื่อตัวแปร

bonus

กำหนดชนิดของข้อมูล

จำนวนจริงทศนิยม 2 ตำแหน่ง

ข้อมูลที่ไม่สมบูรณ์ให้เว้นว่าง

คำแนะนำสำหรับเจ้าหน้าที่พิมพ์ข้อมูล

1. เลขประจำตัว พิมพ์ข้อมูลตามค่าจริงจากแบบสอบถาม

2. เพศ

☐ ชาย

พิมพ์ข้อมูลเป็นเลข 1

☐ หญิง

พิมพ์ข้อมูลเป็นเลข 2

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้พิมพ์ข้อมูลเป็นเลข 9

3. อายุ พิมพ์ข้อมูลตามค่าจริงจากแบบสอบถาม

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้เว้นว่าง

4. ระดับการศึกษา ☐ ต่ำกว่าระดับปริญญาตรี พิมพ์ข้อมูลเป็นเลข 1☐ จบระดับปริญญาตรี พิมพ์ข้อมูลเป็นเลข 2☐ จบระดับปริญญาโท พิมพ์ข้อมูลเป็นเลข 3☐ จบระดับปริญญาเอก พิมพ์ข้อมูลเป็นเลข 4

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้พิมพ์ข้อมูลเป็นเลข 9

5. สถานภาพ ☐ โสด พิมพ์ข้อมูลเป็นเลข 1☐ แต่งงานแล้ว พิมพ์ข้อมูลเป็นเลข 2☐ เป็นหม้าย พิมพ์ข้อมูลเป็นเลข 3☐ หย่าร้าง พิมพ์ข้อมูลเป็นเลข 4

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้พิมพ์ข้อมูลเป็นเลข 9

6. เงินเดือน พิมพ์ข้อมูลตามค่าจริงจากแบบสอบถาม

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้เว้นว่าง

7. ระดับคะแนน พิมพ์ข้อมูลตามค่าจริงจากแบบสอบถาม

8. เงินตอบแทนประจำปี พิมพ์ข้อมูลตามค่าจริงจากแบบสอบถาม

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้เว้นว่าง

ตัวอย่างแบบสอบถามที่กรอกข้อมูลแล้ว

แบบสอบถามข้อมูลพนักงาน

สำหรับเจ้าหน้าที่กรอกข้อมูล

1. เลขประจำตัว 1

2. เพศ ☒ ชาย ☐ หญิง

1

3. อายุ 37 ปี

4. ระดับการศึกษา ☐ ต่ำกว่าระดับปริญญาตรี ☒ จบระดับปริญญาตรี

2

☐ จบระดับปริญญาโท ☐ จบระดับปริญญาเอก

5. สถานภาพ ☐ โสด ☐ แต่งงานแล้ว

4

☐ เป็นหม้าย ☒ หย่าร้าง

6. เงินเดือน 5500 บาท

7. ระดับคะแนน 3.78

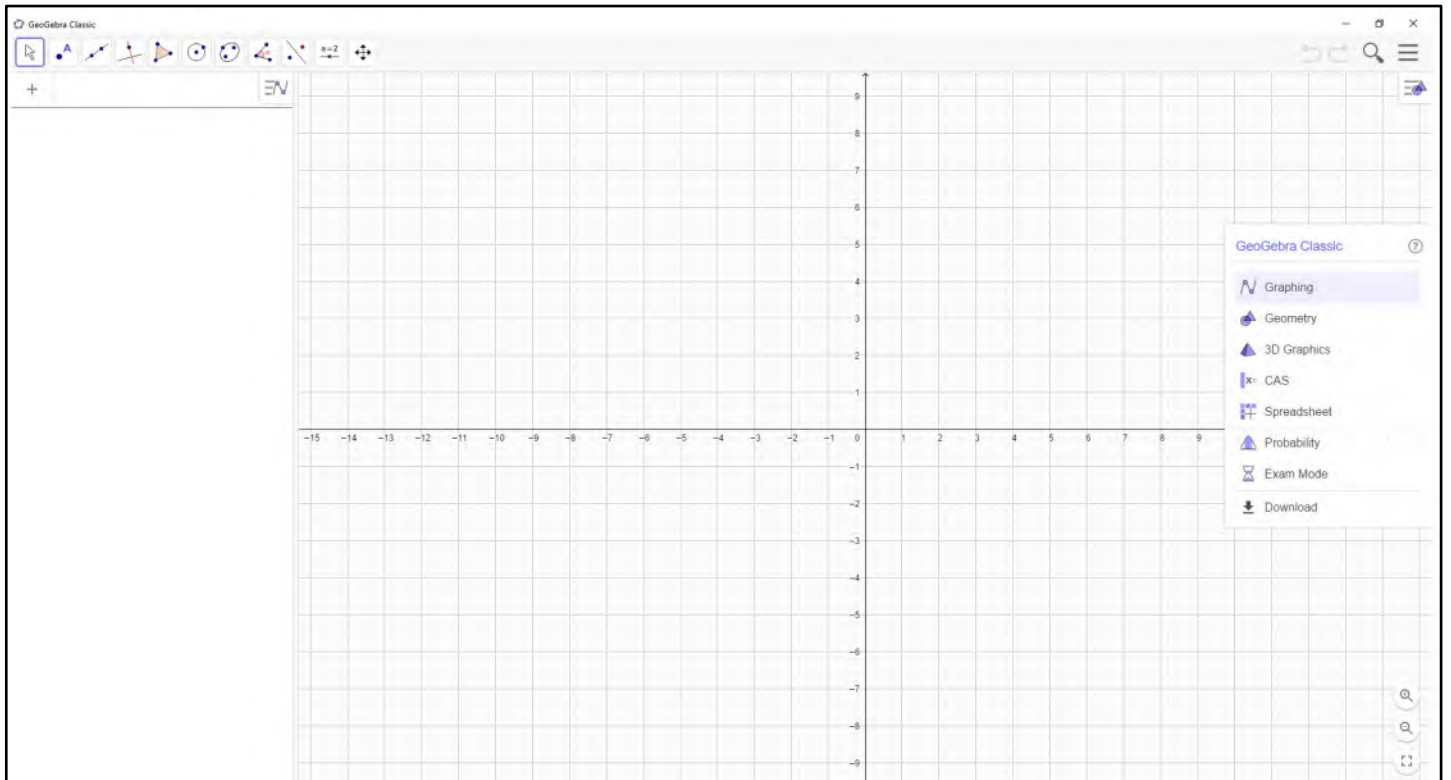
8. เงินตอบแทนประจำปี 11000 บาท

ข้อมูลของพนักงาน 25 คนเป็นดังนี้

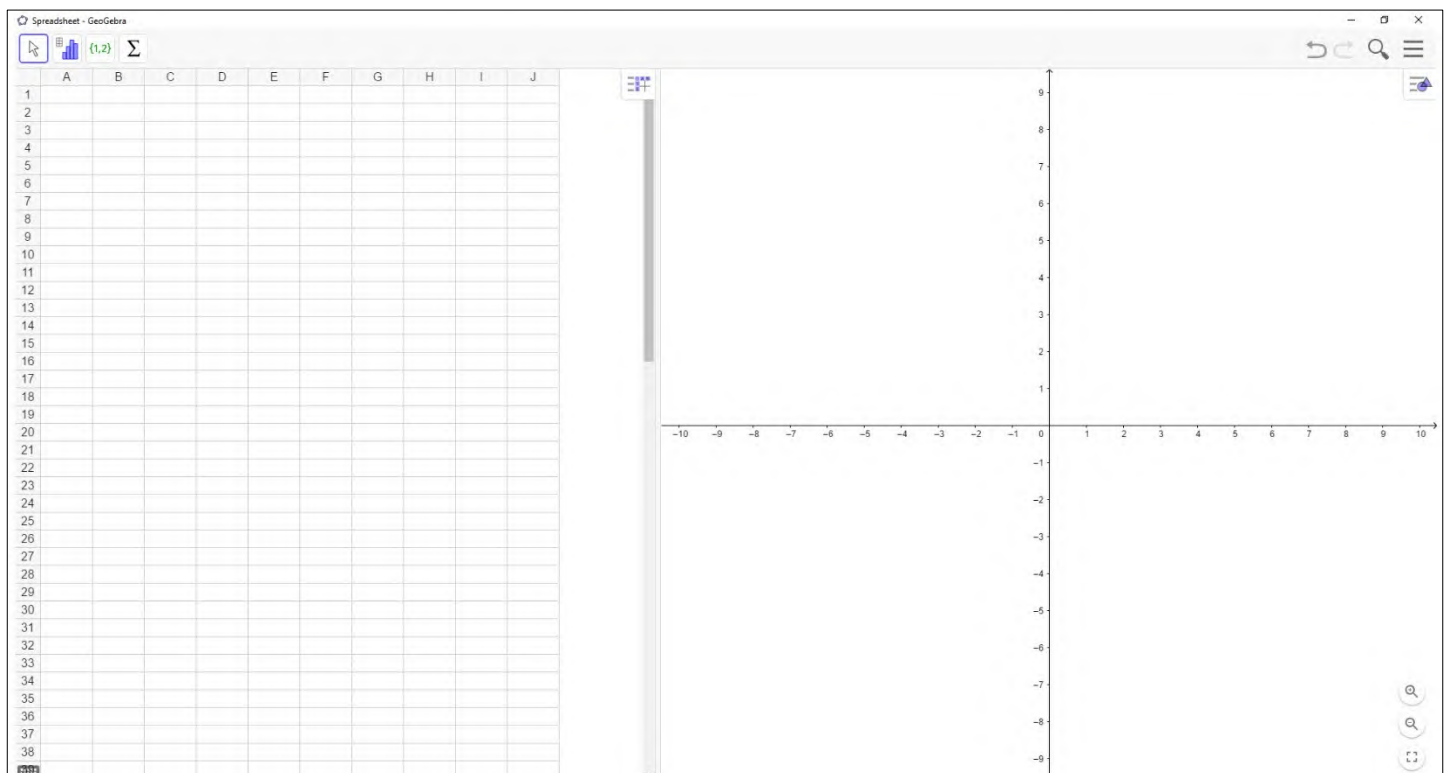
id	sex	age	educ	status	income	grade	bonus
1	1	37	2	4	5500	3.78	11000
2	2	29	3	1	4100	3.89	12300
3	2	48	1	2	5400	3.67	21600
4	1		1	2		2.78	19998
5	2	33	2	9		3.00	29997
6	2	45	3	4	8300	3.45	16600
7	2	38	1	4	7700	3.89	7700
8	2	23	3	1	3900	3.67	11700
9	1	34	2	4	4500	2.56	9000
10	1	50	2	2	6700	2.69	6700
11	2	43	2	2	4700	3.56	18800
12	2	37	3	2	3900	3.00	3900
13	1	24	2	1	3300	2.45	9900
14	1	46	2	2	4900	2.45	14700
15	1	32	1	1	4000	3.87	8000
16	1	42	2	3	6600	3.67	13200
17	1	38	4	2	8000	3.23	32000
18	2	41	2	3	7000	3.45	21000
19	2		1	9	2000	3.21	2000
20	1	54	2	2	7400	3.00	22200
21	2	32	3	9	6200	2.56	24800
22	1	43	1	2	4700	2.45	18800
23	2	22	1	1	3400	3.78	3400
24	1	40	2	2	5900	2.67	17700
25	1	37	4	9	7500	3.45	22500

การสร้างแฟ้มข้อมูลใน Geogebra

ขั้นที่ 1 เปิดโปรแกรม Geogebra



ขั้นที่ 2 เลือก Spreadsheet คลิกที่ Spreadsheet แล้วจะได้หน้าจอดังนี้



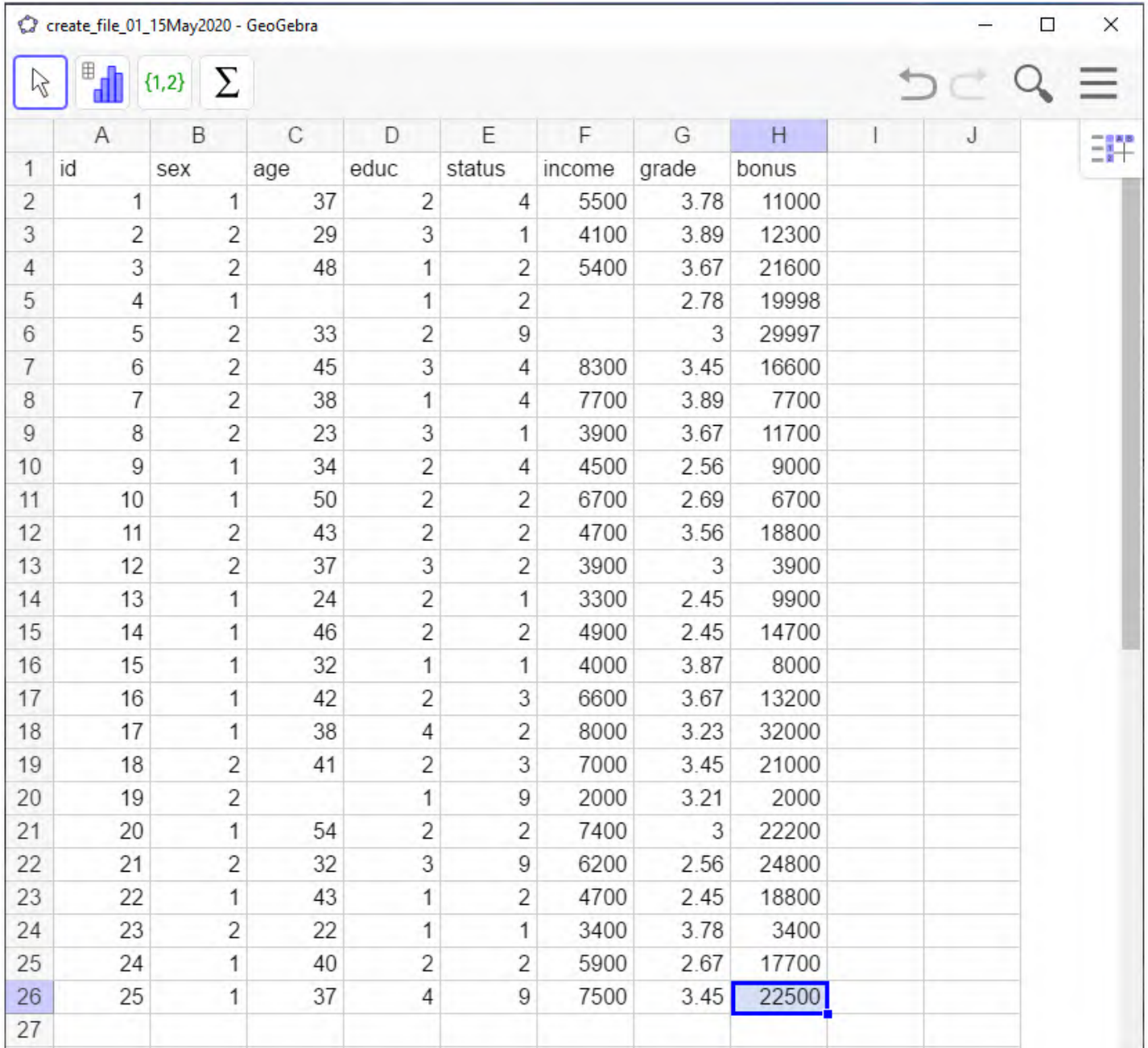
ขั้นที่ 3 พิมพ์ข้อมูล

1. ตัวแปร id, sex, age, educ, status, income, grade, bonus

ให้อยู่ที่ Column A, B, C, D, E, F, G, H ของ Geogebra ตามลำดับ

2. พิมพ์ชื่อตัวแปร id, sex, age, educ, status, income, grade, bonus ใน Row1

พิมพ์ข้อมูลเข้าสู่ Geogebra จะได้ผลบนจอภาพดังนี้



	A	B	C	D	E	F	G	H	I	J
1	id	sex	age	educ	status	income	grade	bonus		
2	1	1	37	2	4	5500	3.78	11000		
3	2	2	29	3	1	4100	3.89	12300		
4	3	2	48	1	2	5400	3.67	21600		
5	4	1		1	2		2.78	19998		
6	5	2	33	2	9		3	29997		
7	6	2	45	3	4	8300	3.45	16600		
8	7	2	38	1	4	7700	3.89	7700		
9	8	2	23	3	1	3900	3.67	11700		
10	9	1	34	2	4	4500	2.56	9000		
11	10	1	50	2	2	6700	2.69	6700		
12	11	2	43	2	2	4700	3.56	18800		
13	12	2	37	3	2	3900	3	3900		
14	13	1	24	2	1	3300	2.45	9900		
15	14	1	46	2	2	4900	2.45	14700		
16	15	1	32	1	1	4000	3.87	8000		
17	16	1	42	2	3	6600	3.67	13200		
18	17	1	38	4	2	8000	3.23	32000		
19	18	2	41	2	3	7000	3.45	21000		
20	19	2		1	9	2000	3.21	2000		
21	20	1	54	2	2	7400	3	22200		
22	21	2	32	3	9	6200	2.56	24800		
23	22	1	43	1	2	4700	2.45	18800		
24	23	2	22	1	1	3400	3.78	3400		
25	24	1	40	2	2	5900	2.67	17700		
26	25	1	37	4	9	7500	3.45	22500		
27										

หมายเหตุ: แฟ้มข้อมูลชื่อ create_file_01_15May2020.ggb

1.3 การหาค่าสถิติเบื้องต้น

การหาค่าสถิติเบื้องต้นของตัวแปร income โดยใช้โปรแกรม Geogebra

ขั้นที่ 1 คลิกที่ column ตัวแปร income

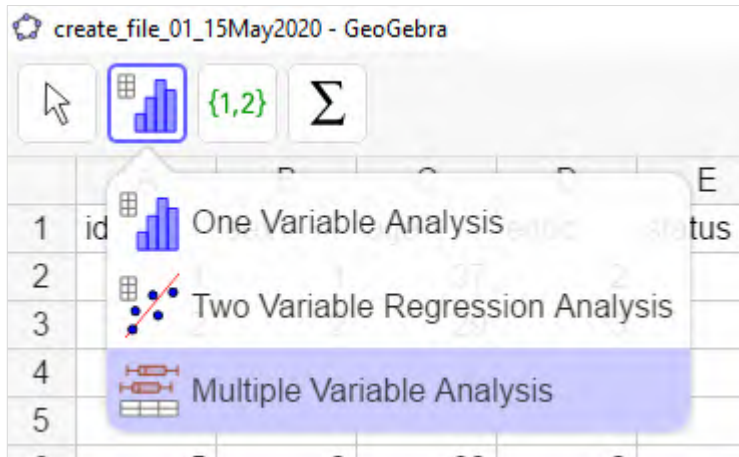
create_file_01_15May2020 - GeoGebra



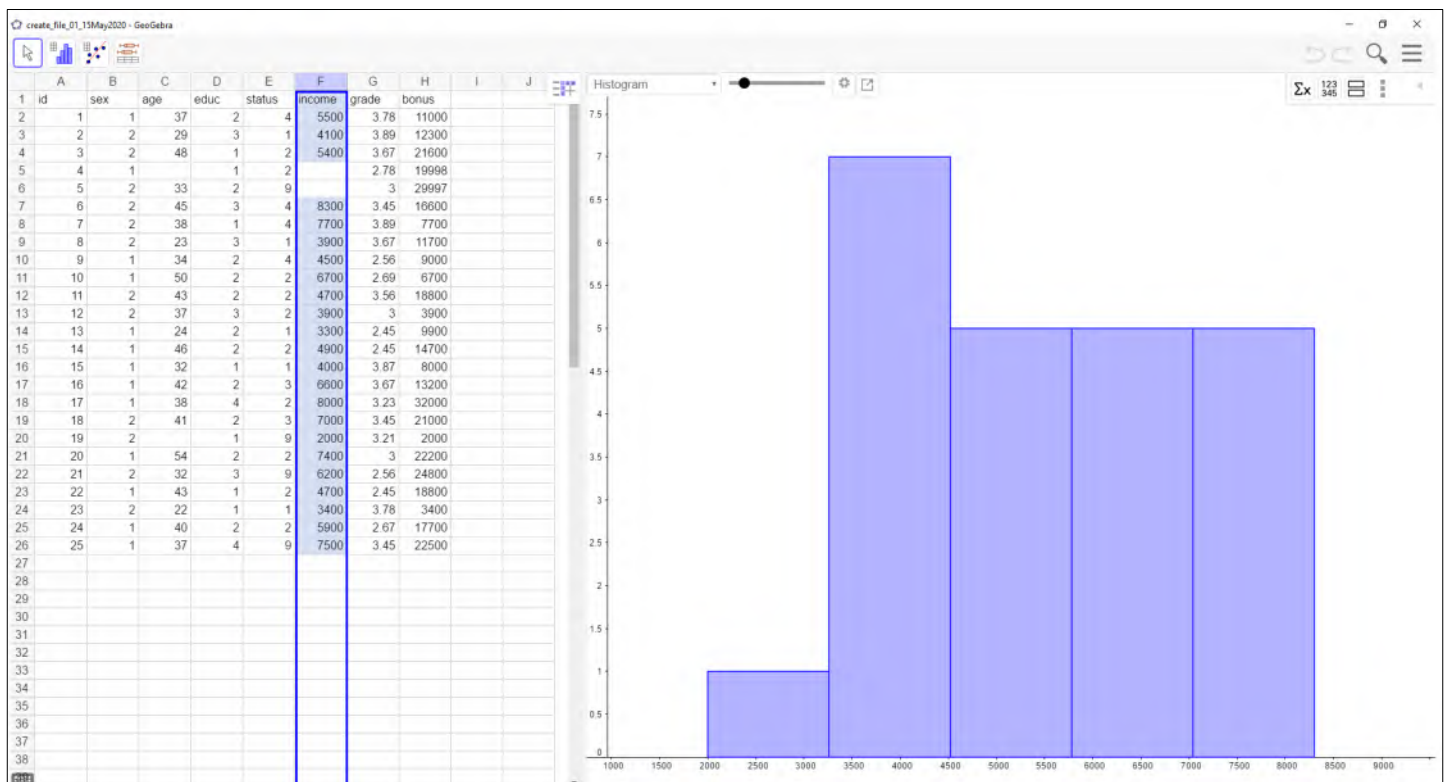
{1,2}


	A	B	C	D	E	F	G	H
1	id	sex	age	educ	status	income	grade	bonus
2	1	1	37	2	4	5500	3.78	11000
3	2	2	29	3	1	4100	3.89	12300
4	3	2	48	1	2	5400	3.67	21600
5	4	1		1	2		2.78	19998
6	5	2	33	2	9		3	29997
7	6	2	45	3	4	8300	3.45	16600
8	7	2	38	1	4	7700	3.89	7700
9	8	2	23	3	1	3900	3.67	11700
10	9	1	34	2	4	4500	2.56	9000
11	10	1	50	2	2	6700	2.69	6700
12	11	2	43	2	2	4700	3.56	18800
13	12	2	37	3	2	3900	3	3900
14	13	1	24	2	1	3300	2.45	9900
15	14	1	46	2	2	4900	2.45	14700
16	15	1	32	1	1	4000	3.87	8000
17	16	1	42	2	3	6600	3.67	13200
18	17	1	38	4	2	8000	3.23	32000
19	18	2	41	2	3	7000	3.45	21000
20	19	2		1	9	2000	3.21	2000
21	20	1	54	2	2	7400	3	22200
22	21	2	32	3	9	6200	2.56	24800
23	22	1	43	1	2	4700	2.45	18800
24	23	2	22	1	1	3400	3.78	3400
25	24	1	40	2	2	5900	2.67	17700
26	25	1	37	4	9	7500	3.45	22500

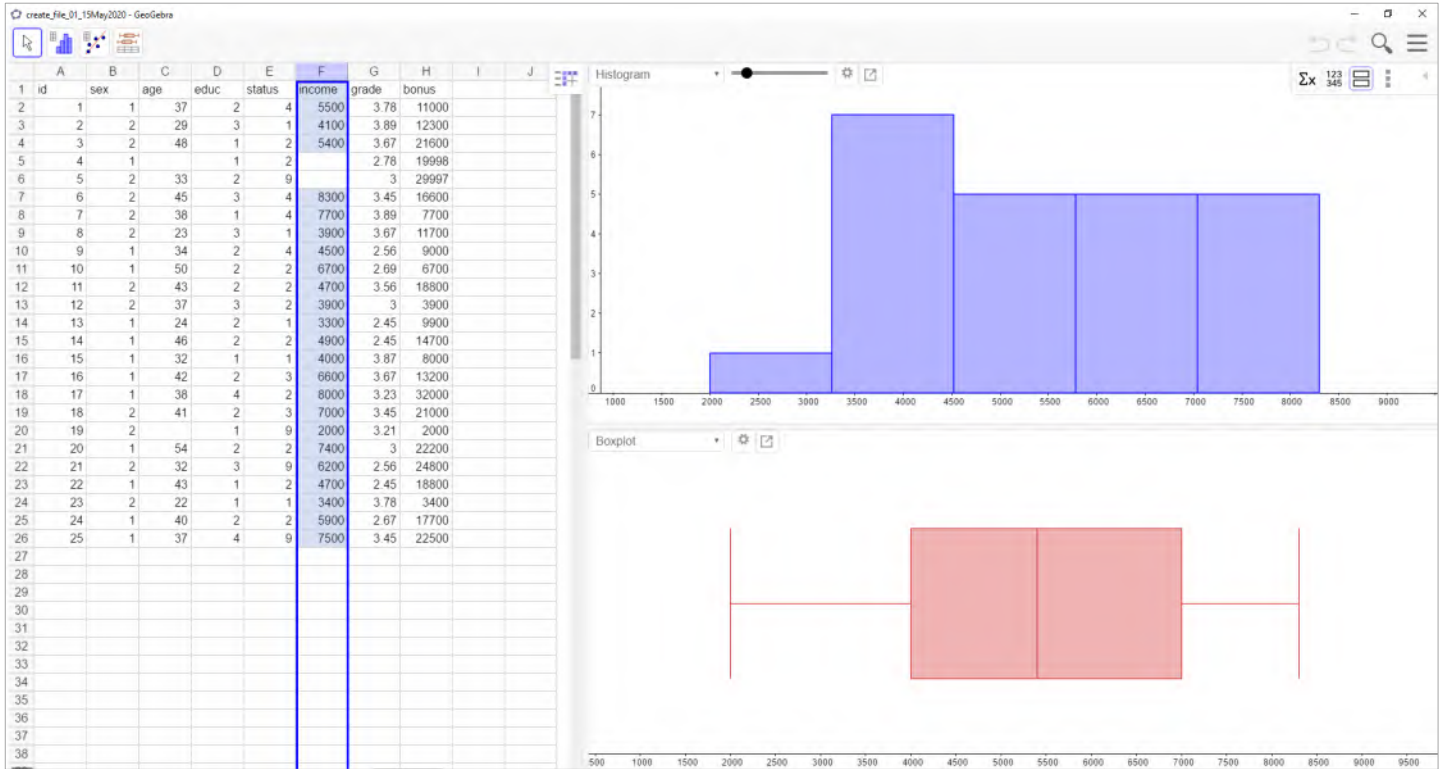
ขั้นที่ 2 คลิกที่  เมื่อคลิกแล้วจะได้เมนูย่อย



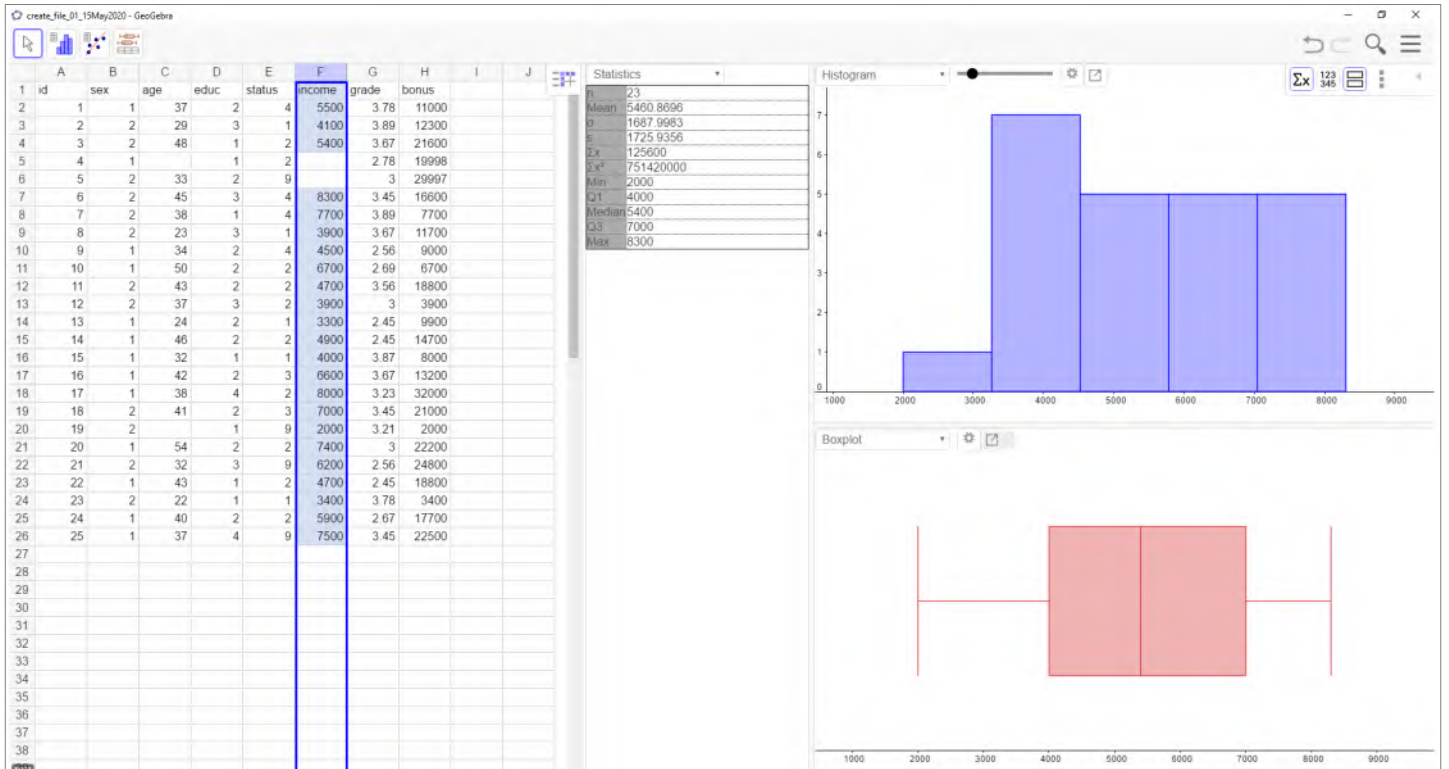
ขั้นที่ 3 คลิกที่  One Variable Analysis เมื่อคลิกแล้วจะได้



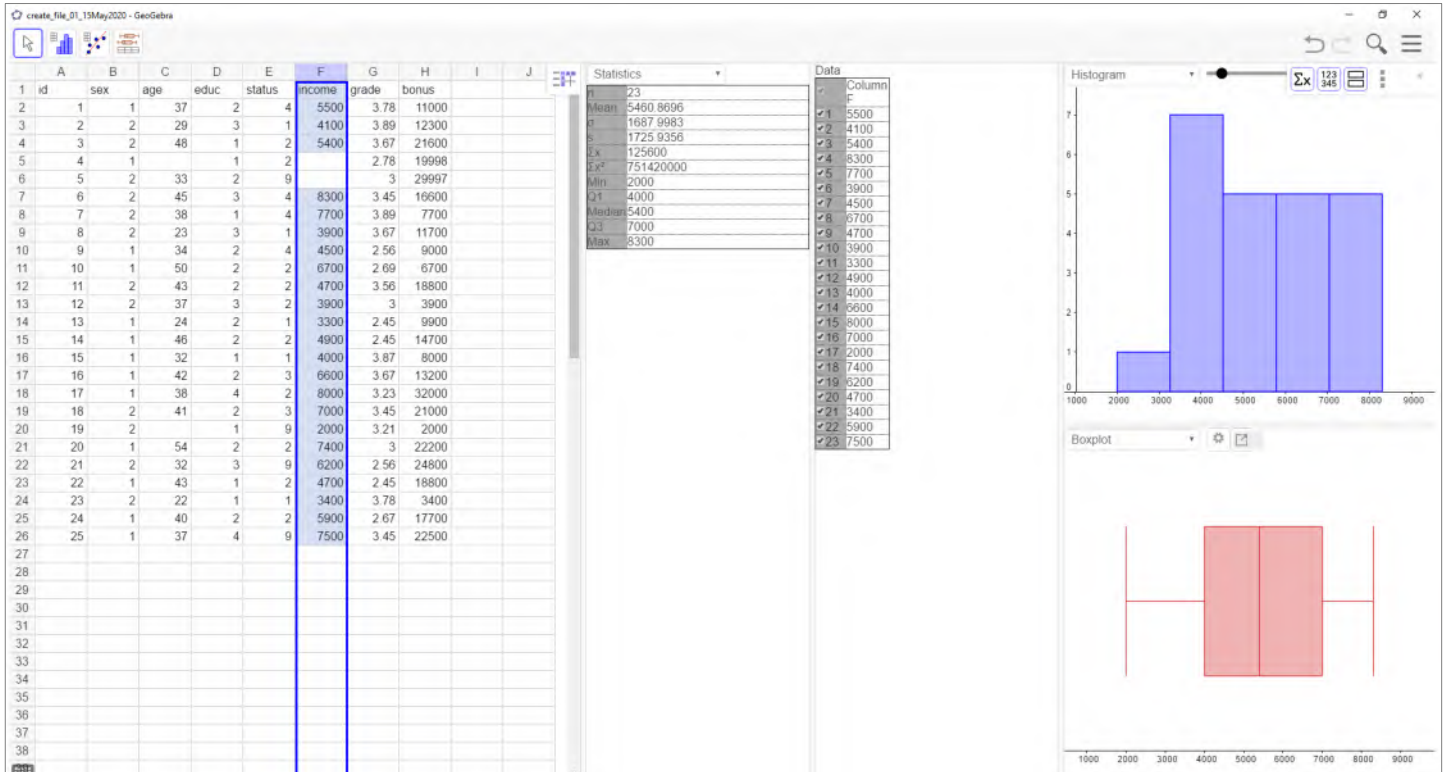
ขั้นที่ 4 คลิกที่  เมื่อคลิกแล้วจะได้



ขั้นที่ 5 คลิกที่ Σx เมื่อคลิกแล้วจะได้



ขั้นที่ 6 คลิกที่ **123**
345 เมื่อคลิกแล้วจะได้



ข้อมูลจริงที่นำมาวิเคราะห์ข้อมูล 23 ตัว จาก column F

หมายเหตุ หากต้องการยกเลิกการแสดงตาราง Data ให้คลิกที่ **123**
345

หมายเหตุ ผลการคำนวณเป็นดังนี้

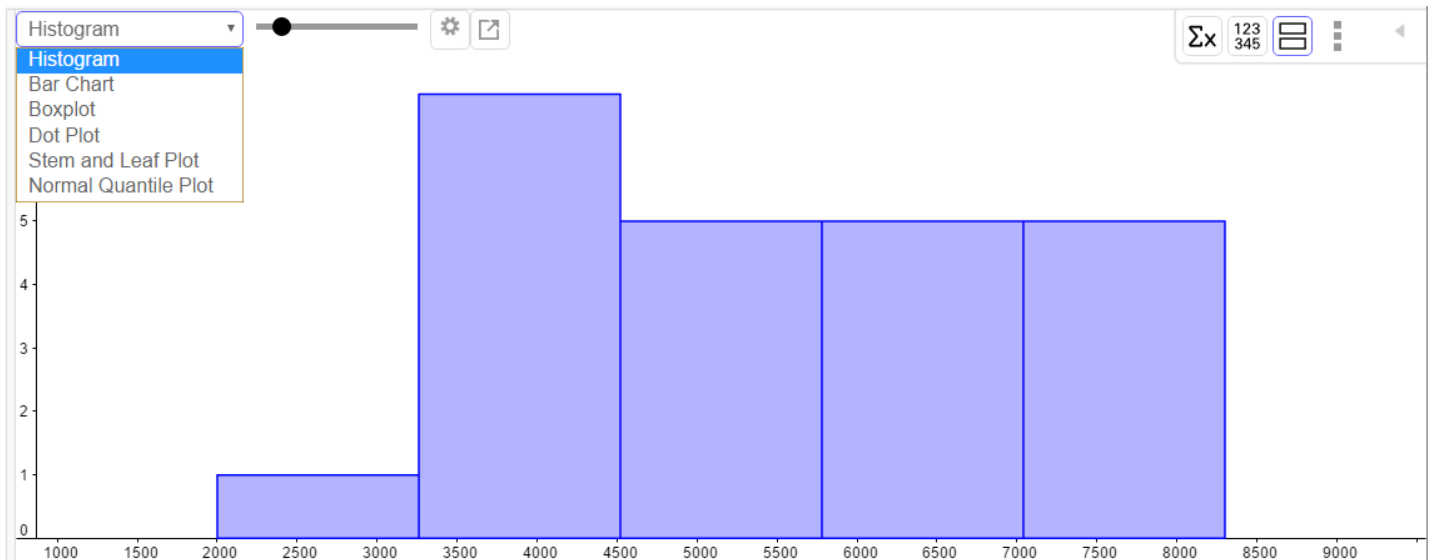
Statistics	
n	23
Mean	5460.8696
σ	1687.9983
s	1725.9356
Σx	125600
Σx^2	751420000
Min	2000
Q1	4000
Median	5400
Q3	7000
Max	8300

1. จำนวนข้อมูล $n = 23$ (มีพนักงานไม่กรอกรายได้ 2 คน)
2. ค่าเฉลี่ยของตัวอย่าง $\bar{x} = 5460.8696$ (คิดจากพนักงาน 23 คน $\bar{x} = \frac{125600}{23}$)
3. ค่าส่วนเบี่ยงเบนมาตรฐาน (สูตรประชากร) $\sigma = 1687.9983$
3. ค่าส่วนเบี่ยงเบนมาตรฐาน (สูตรตัวอย่าง) $s = 1725.9356$
4. ผลรวมรายได้ $\sum x = 125600$
5. $\sum x^2 = 751420000$
6. ค่าต่ำสุด Minimum = 2000
7. ควอไทล์ที่ 1 คือ 4000
8. มัธยฐาน = 5400
9. ควอไทล์ที่ 3 คือ 7000
10. ค่าสูงสุด Maximum = 8300

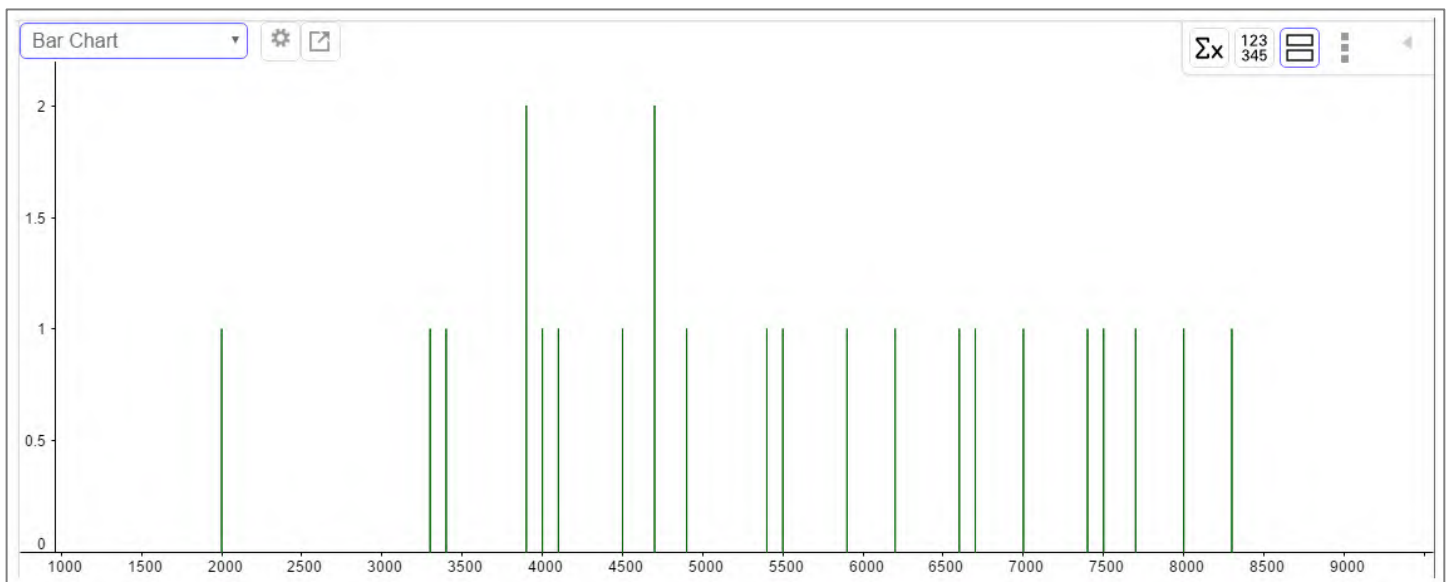
หมายเหตุ หากต้องการยกเลิกการแสดงตาราง Statistics ให้คลิกที่ 

การเขียนกราฟแบบต่าง ๆ

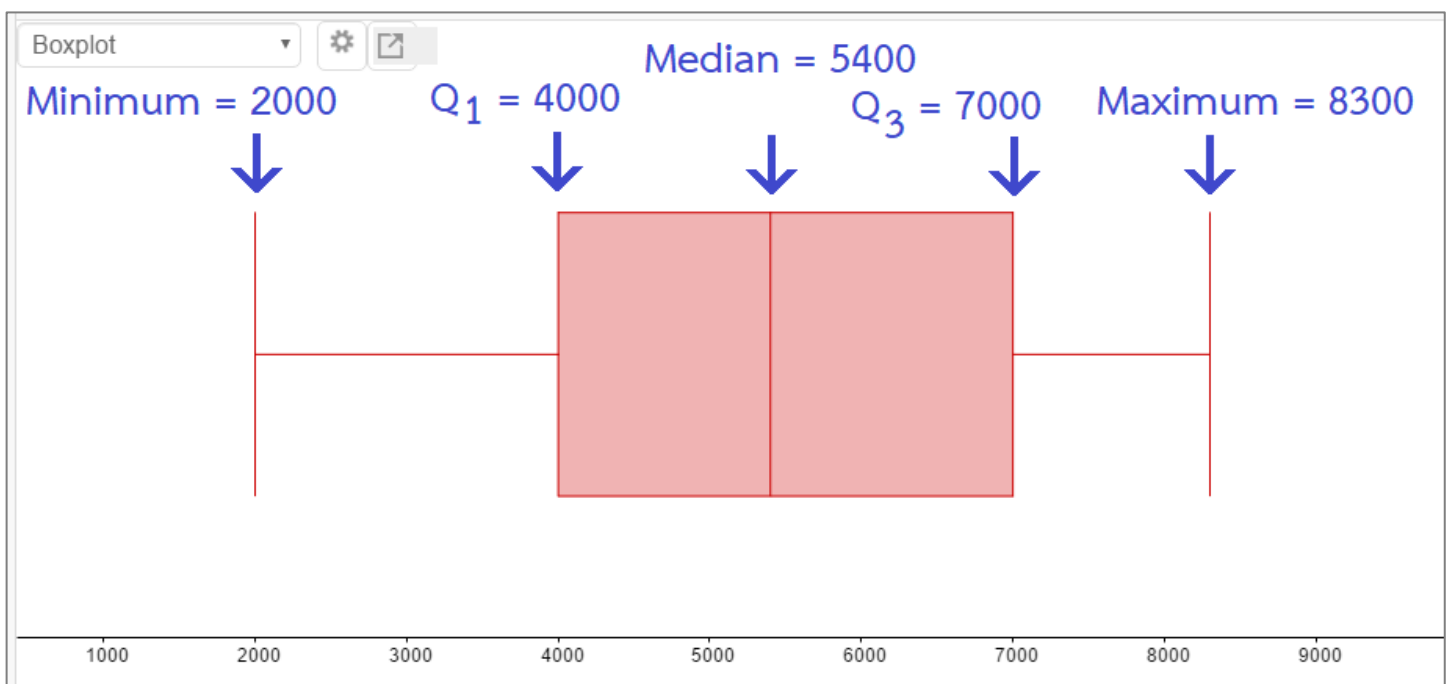
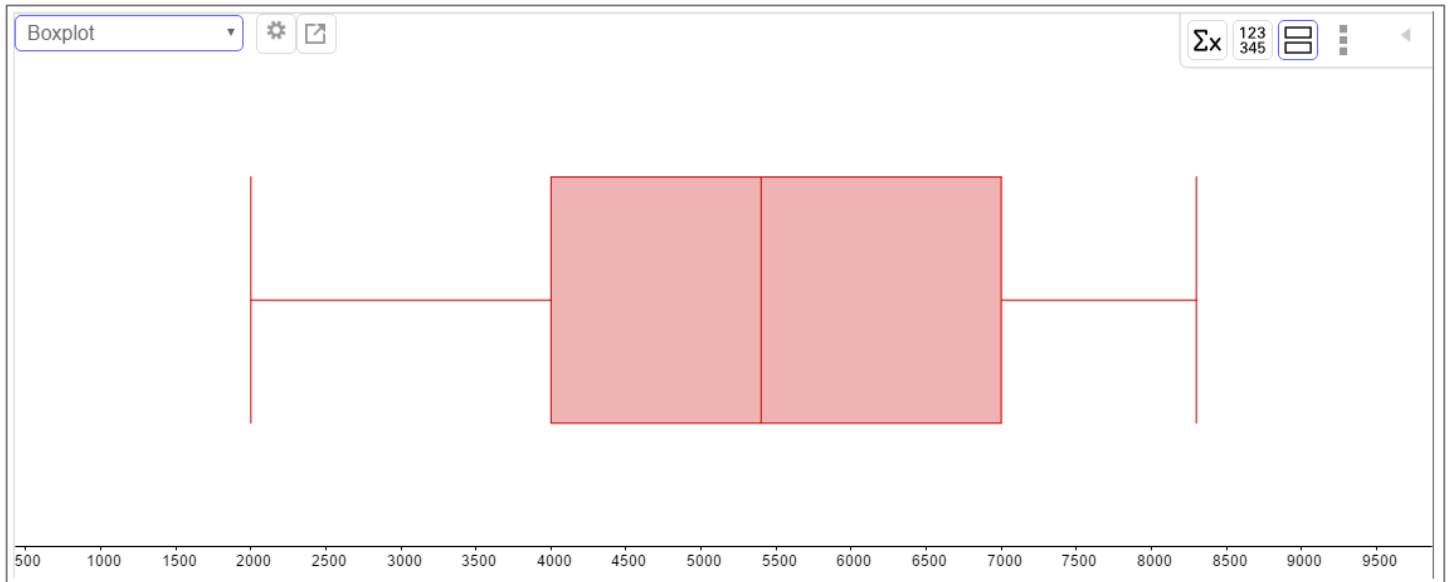
คลิกตรงนี้จะได้เมนูย่อยของ
การเขียนกราฟ



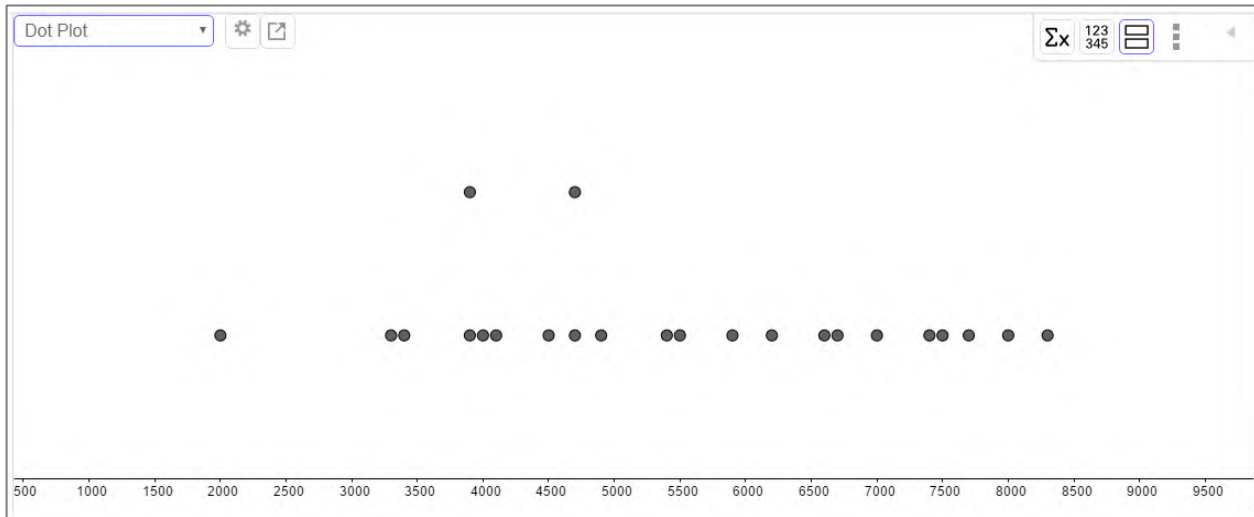
เลือก **Bar Chart** คลิกแล้วจะได้กราฟ



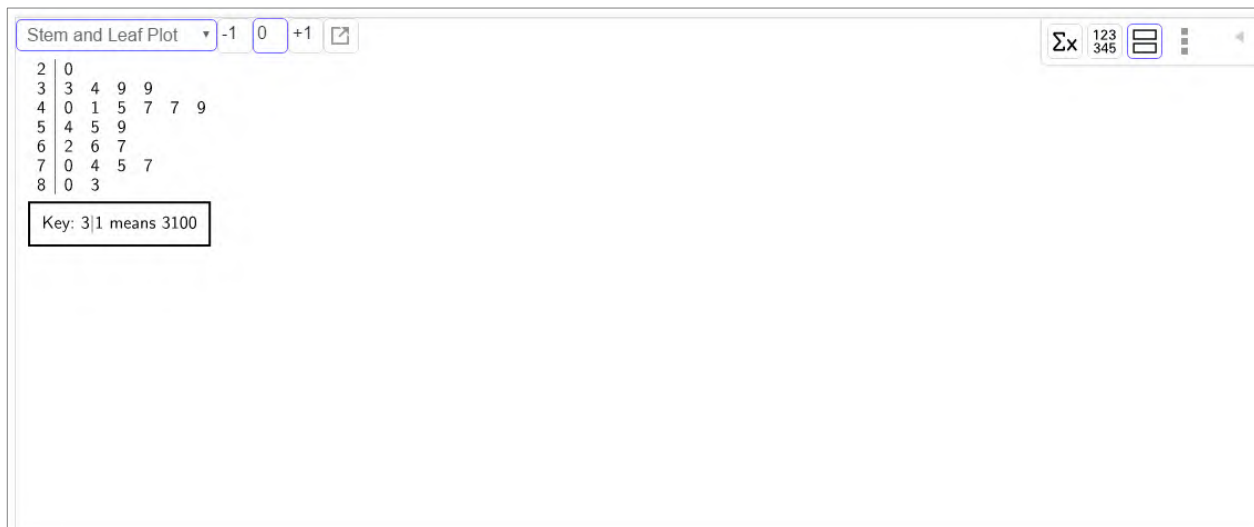
เลือก **Boxplot** คลิกแล้วจะได้กราฟ



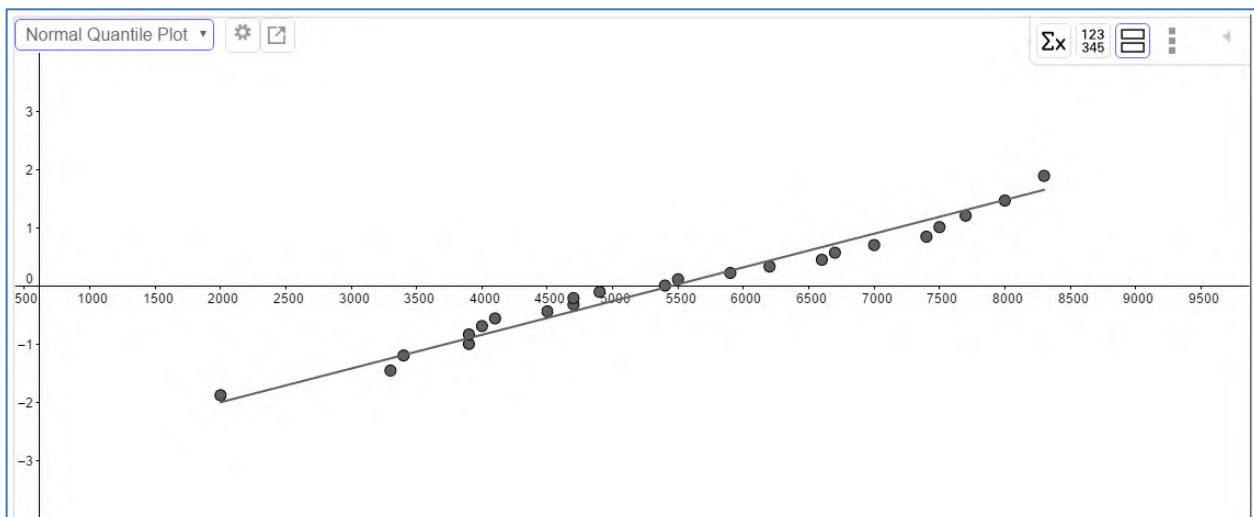
เลือก **Dot Plot** คลิกแล้วจะได้กราฟ



เลือก **Stem and Leaf Plot** คลิกแล้วจะได้กราฟ



เลือก **Normal Quantile Plot** คลิกแล้วจะได้กราฟ



การเพิ่ม/ลด ช่วงของการเขียนกราฟแบบ Histogram



การหาค่าสถิติเบื้องต้นของตัวแปร age

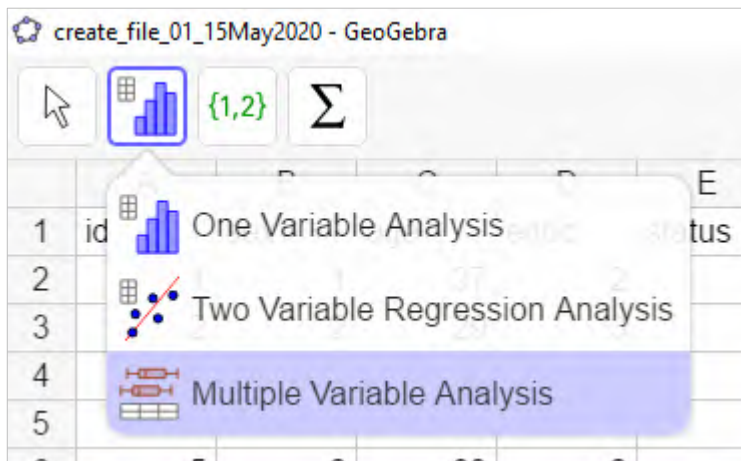
ขั้นที่ 1 คลิกที่ column ตัวแปร age

create_file_01_15May2020 - GeoGebra

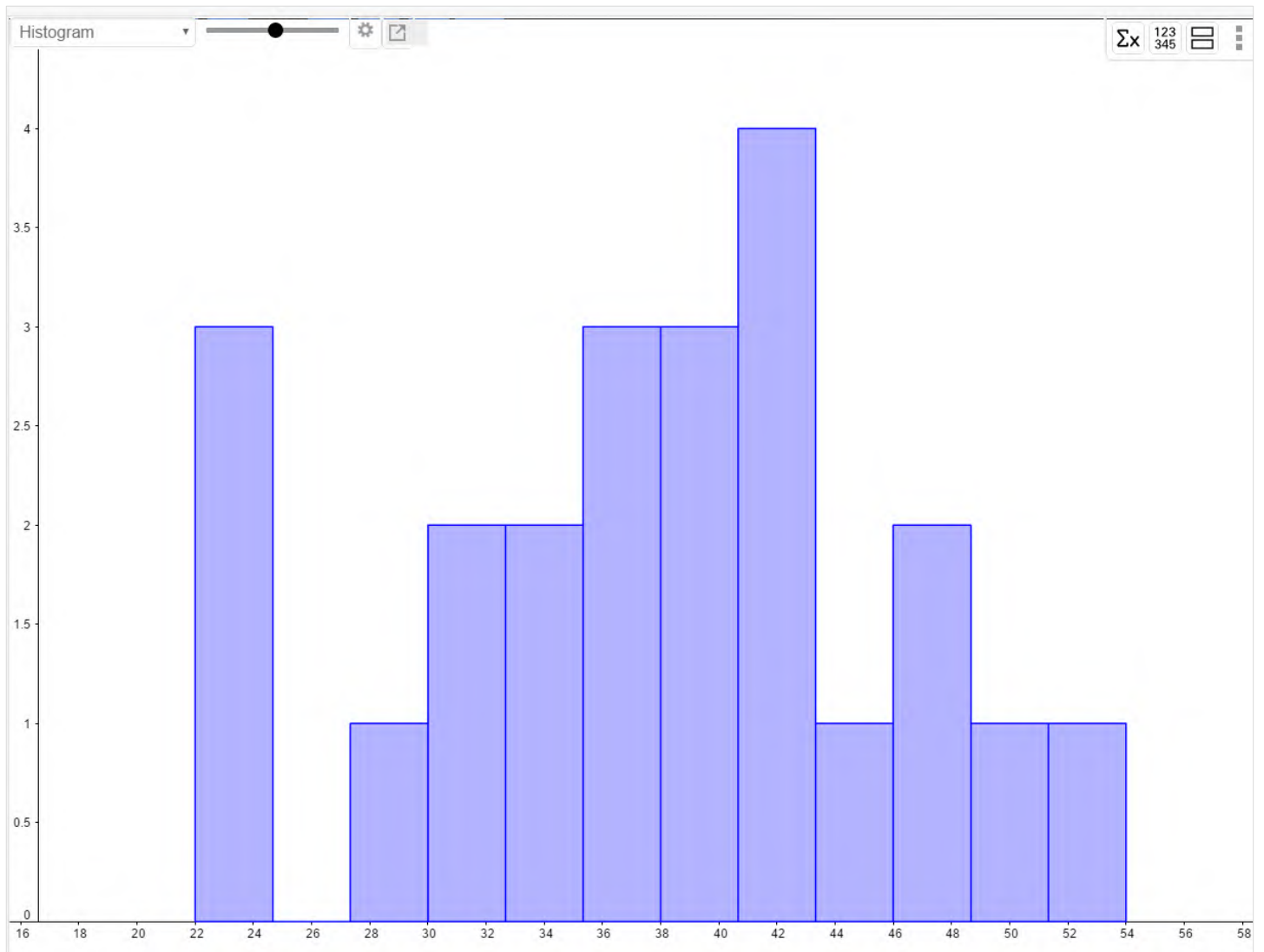


	A	B	C	D	E	F	G	H
1	id	sex	age	educ	status	income	grade	bonus
2	1	1	37	2	4	5500	3.78	11000
3	2	2	29	3	1	4100	3.89	12300
4	3	2	48	1	2	5400	3.67	21600
5	4	1		1	2		2.78	19998
6	5	2	33	2	9		3	29997
7	6	2	45	3	4	8300	3.45	16600
8	7	2	38	1	4	7700	3.89	7700
9	8	2	23	3	1	3900	3.67	11700
10	9	1	34	2	4	4500	2.56	9000
11	10	1	50	2	2	6700	2.69	6700
12	11	2	43	2	2	4700	3.56	18800
13	12	2	37	3	2	3900	3	3900
14	13	1	24	2	1	3300	2.45	9900
15	14	1	46	2	2	4900	2.45	14700
16	15	1	32	1	1	4000	3.87	8000
17	16	1	42	2	3	6600	3.67	13200
18	17	1	38	4	2	8000	3.23	32000
19	18	2	41	2	3	7000	3.45	21000
20	19	2		1	9	2000	3.21	2000
21	20	1	54	2	2	7400	3	22200
22	21	2	32	3	9	6200	2.56	24800
23	22	1	43	1	2	4700	2.45	18800
24	23	2	22	1	1	3400	3.78	3400
25	24	1	40	2	2	5900	2.67	17700
26	25	1	37	4	9	7500	3.45	22500

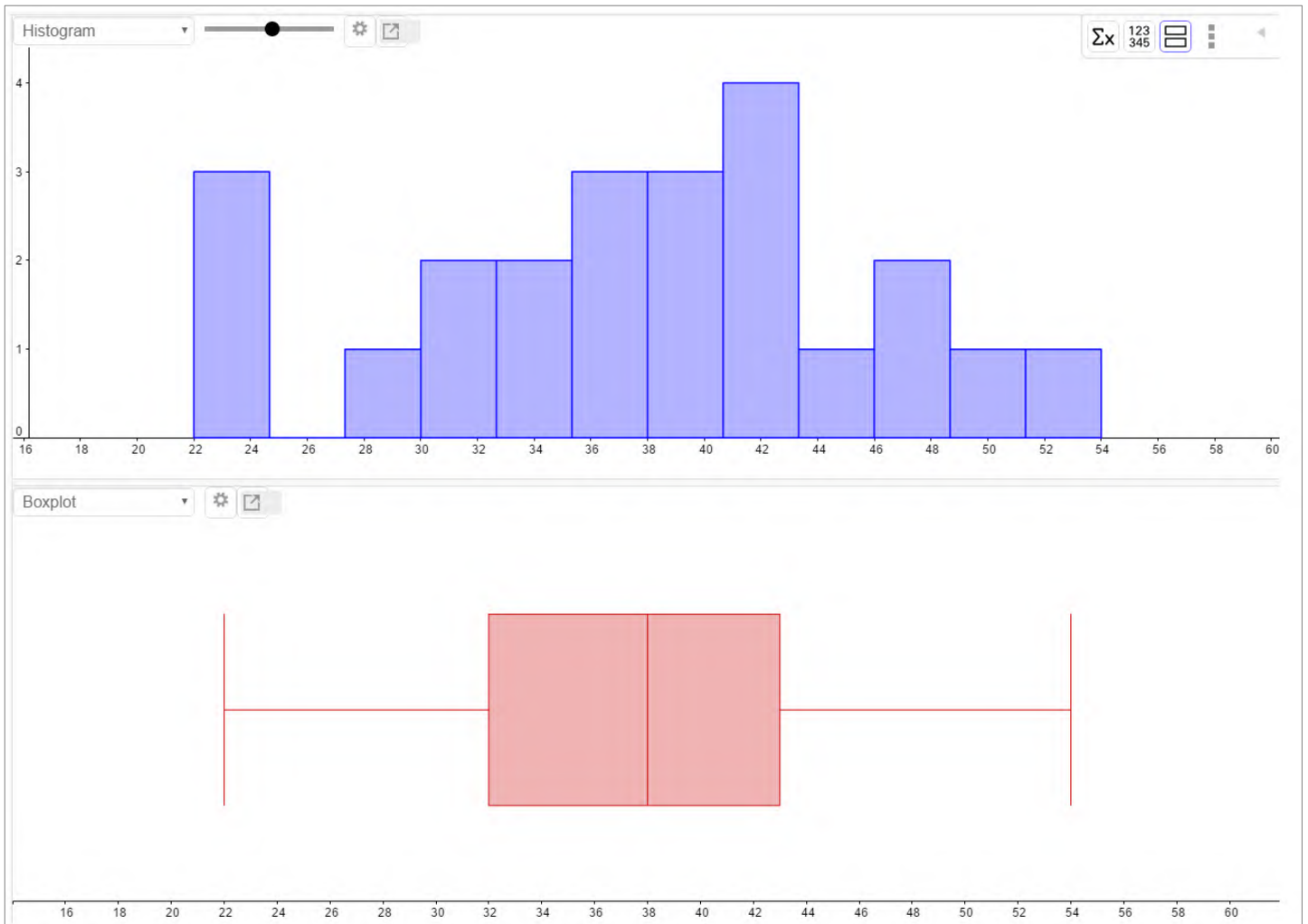
ขั้นที่ 2 คลิกที่  เมื่อคลิกแล้วจะได้เมนูย่อย



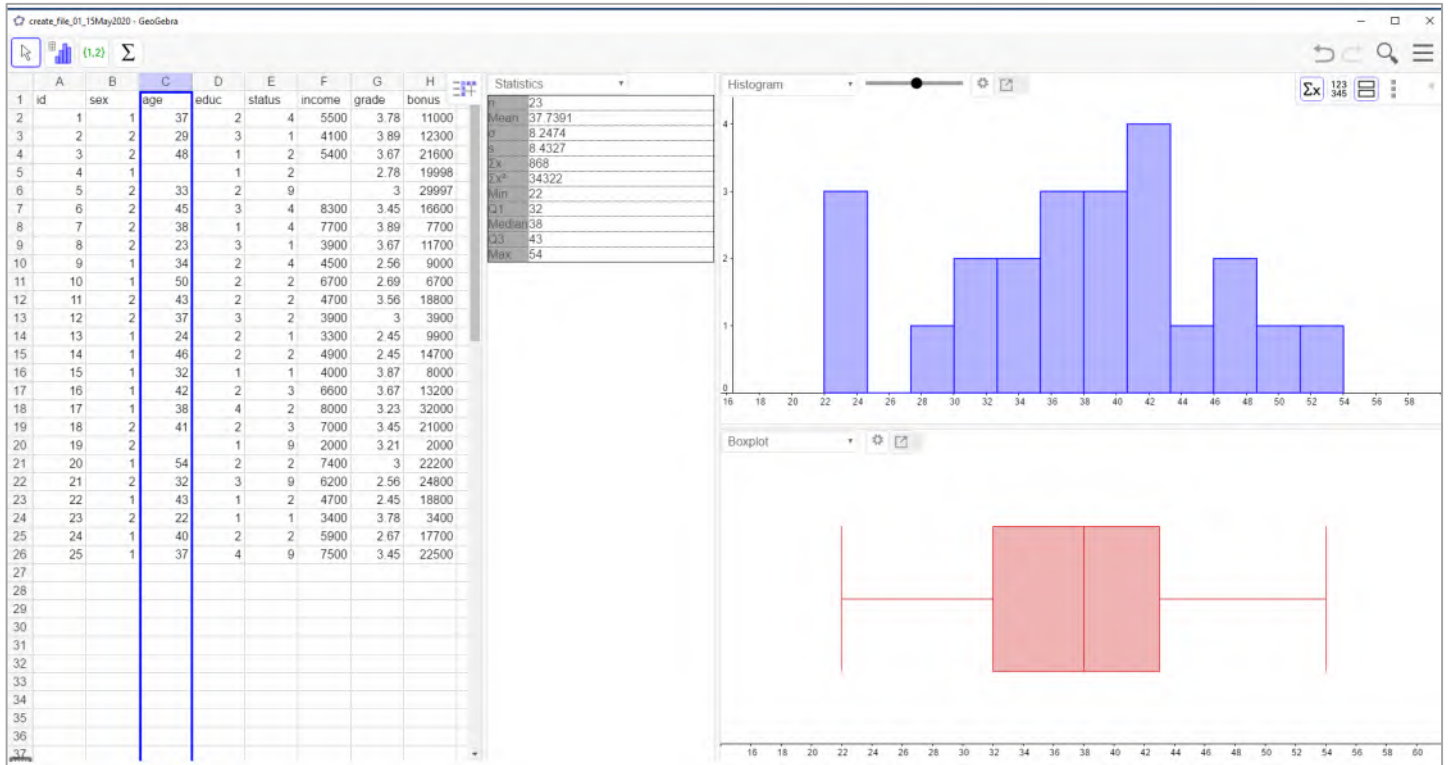
ขั้นที่ 3 คลิกที่  One Variable Analysis เมื่อคลิกแล้วจะได้



ขั้นที่ 4 คลิกที่  เมื่อคลิกแล้วจะได้



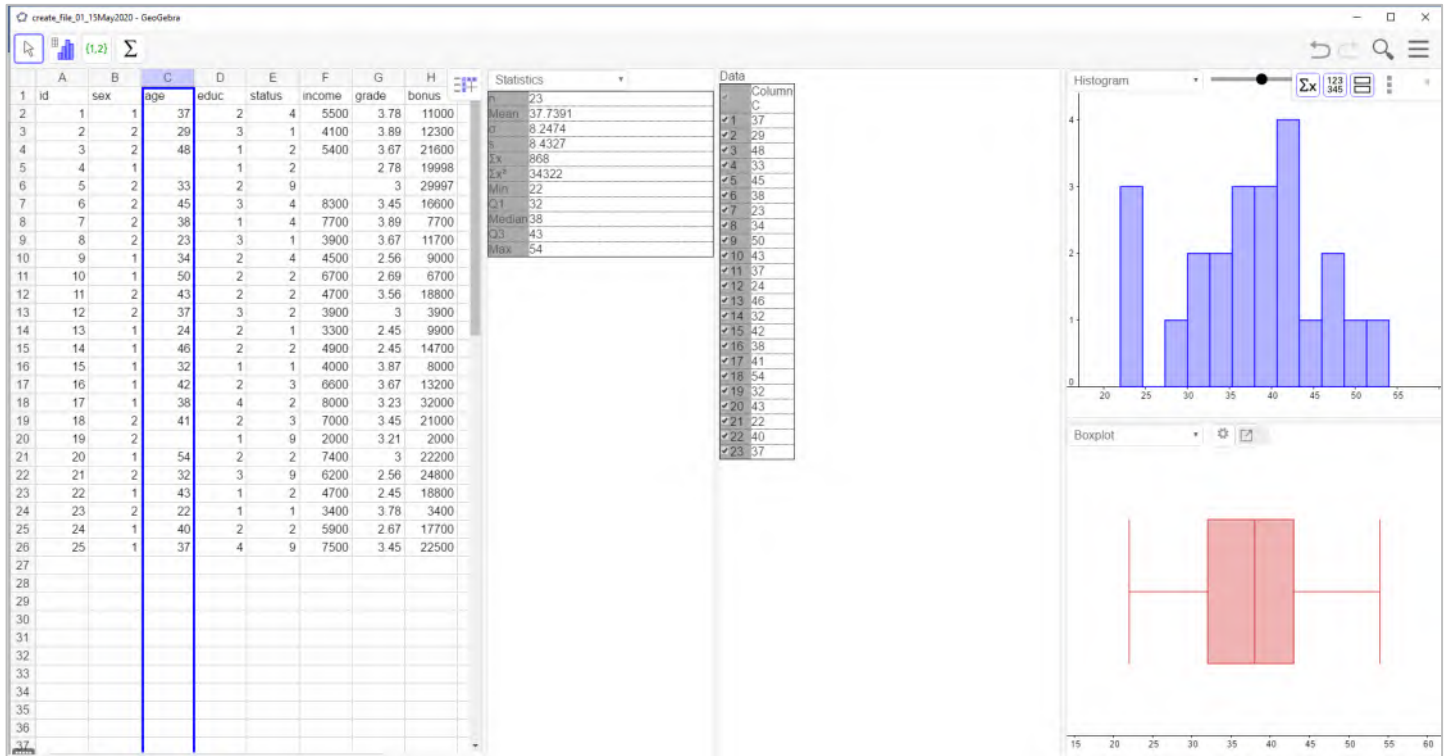
ขั้นที่ 5 คลิกที่ Σx เมื่อคลิกแล้วจะได้



Statistics	
n	23
Mean	37.7391
σ	8.2474
s	8.4327
Σx	868
Σx^2	34322
Min	22
Q1	32
Median	38
Q3	43
Max	54

หมายเหตุ หากต้องการยกเลิกการแสดงตาราง Statistics ให้คลิกที่ Σx

ขั้นที่ 6 คลิกที่ **123**
345 เมื่อคลิกแล้วจะได้



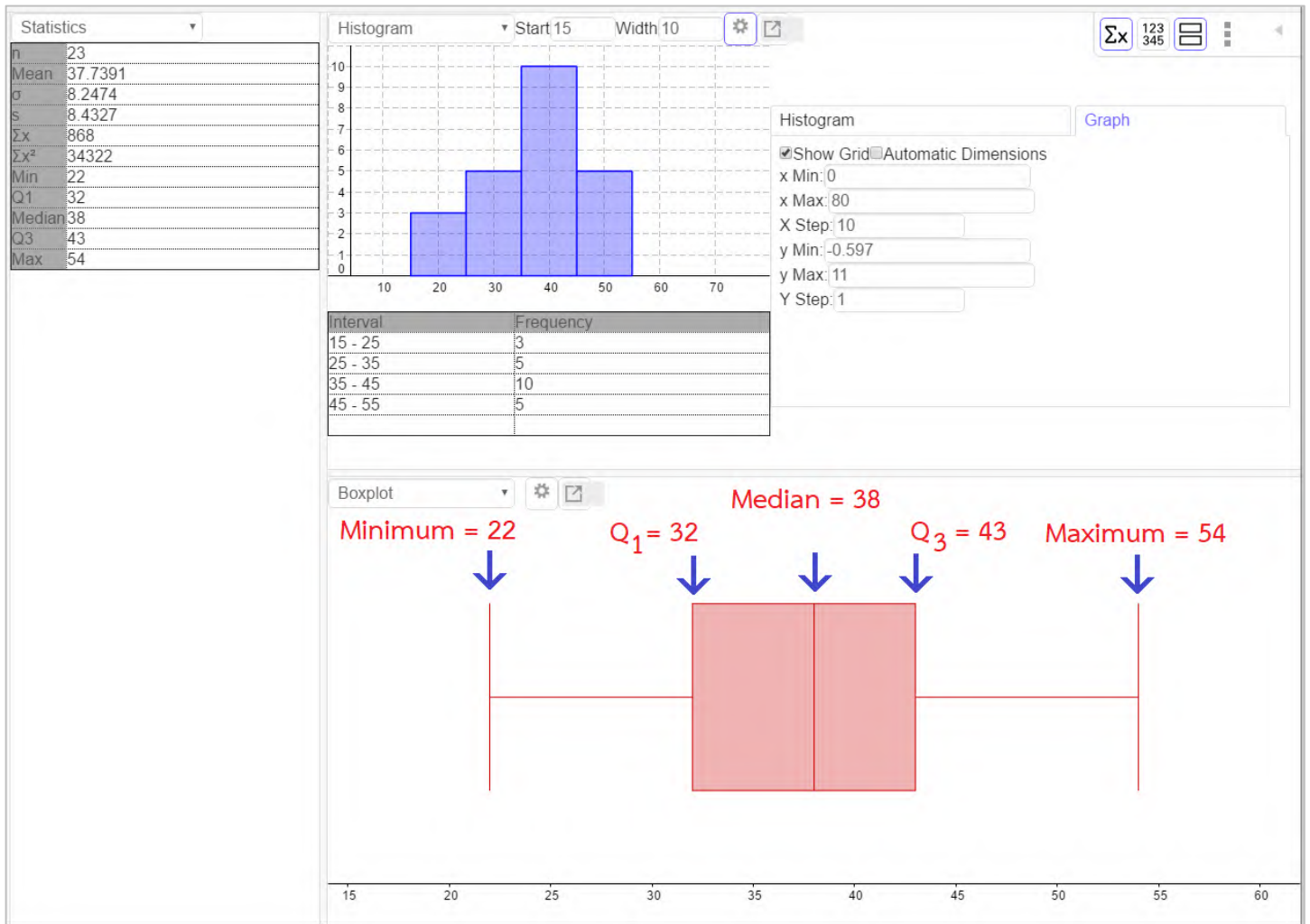
ข้อมูลจริงที่นำมาวิเคราะห์ข้อมูล 23 ตัว จาก column C

หมายเหตุ หากต้องการยกเลิกการแสดงตาราง Data ให้คลิกที่ **123**
345

หมายเหตุ ผลการคำนวณเป็นดังนี้

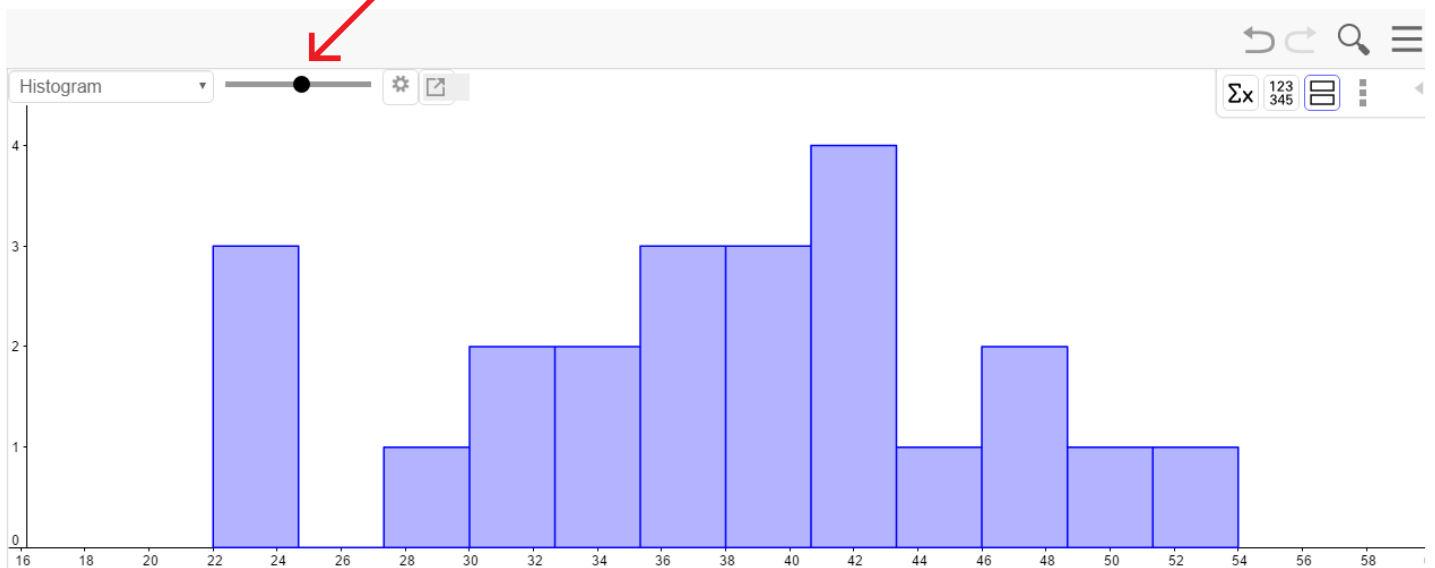
Statistics	
n	23
Mean	37.7391
σ	8.2474
s	8.4327
Σx	868
Σx^2	34322
Min	22
Q1	32
Median	38
Q3	43
Max	54

1. จำนวนข้อมูล $n = 23$ (มีพนักงานไม่กรอกรายได้ 2 คน)
2. ค่าเฉลี่ยของตัวอย่าง $\bar{x} = 37.7391$ (คิดจากพนักงาน 23 คน $\bar{x} = \frac{868}{23}$)
3. ค่าส่วนเบี่ยงเบนมาตรฐาน (สูตรประชากร) $\sigma = 8.2474$
3. ค่าส่วนเบี่ยงเบนมาตรฐาน (สูตรตัวอย่าง) $s = 8.4327$
4. ผลรวมรายได้ $\Sigma x = 868$
5. $\Sigma x^2 = 34322$
6. ค่าต่ำสุด Minimum = 22
7. ควอไทล์ที่ 1 คือ 32
8. มัธยฐาน = 38
9. ควอไทล์ที่ 3 คือ 43
10. ค่าสูงสุด Maximum = 54

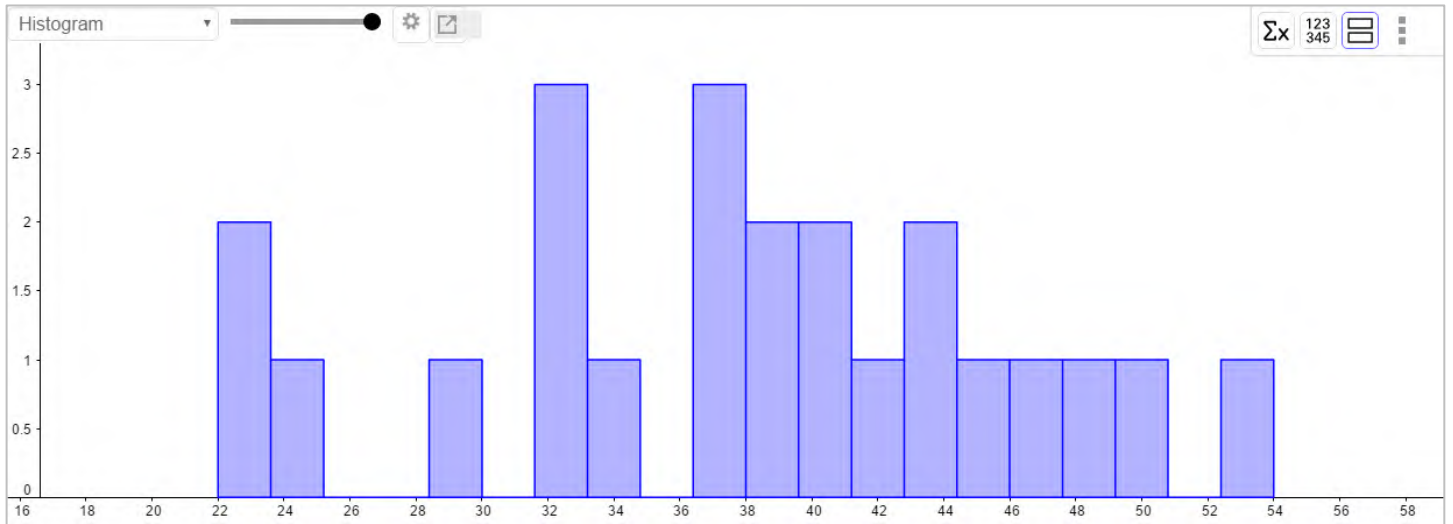


การปรับค่าของการเขียนกราฟแบบ Histogram

เลื่อน ข้าย/ขวา เพื่อ เพิ่ม/ลด จำนวนช่วง
ของกราฟ



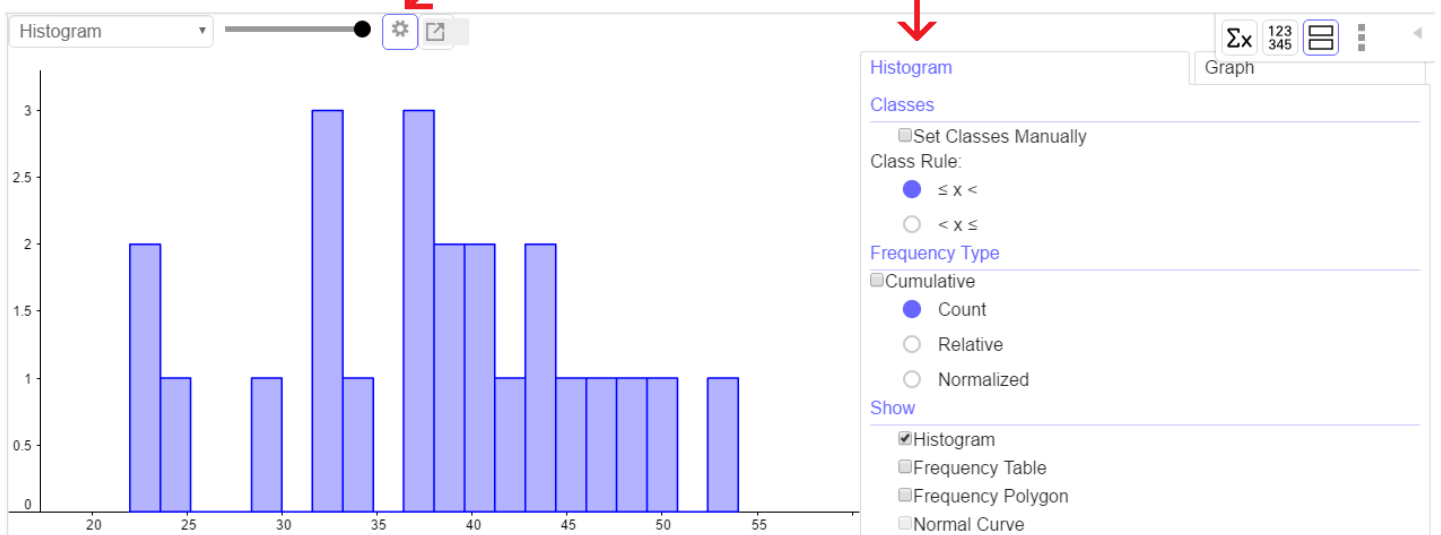
ตัวอย่างเช่น



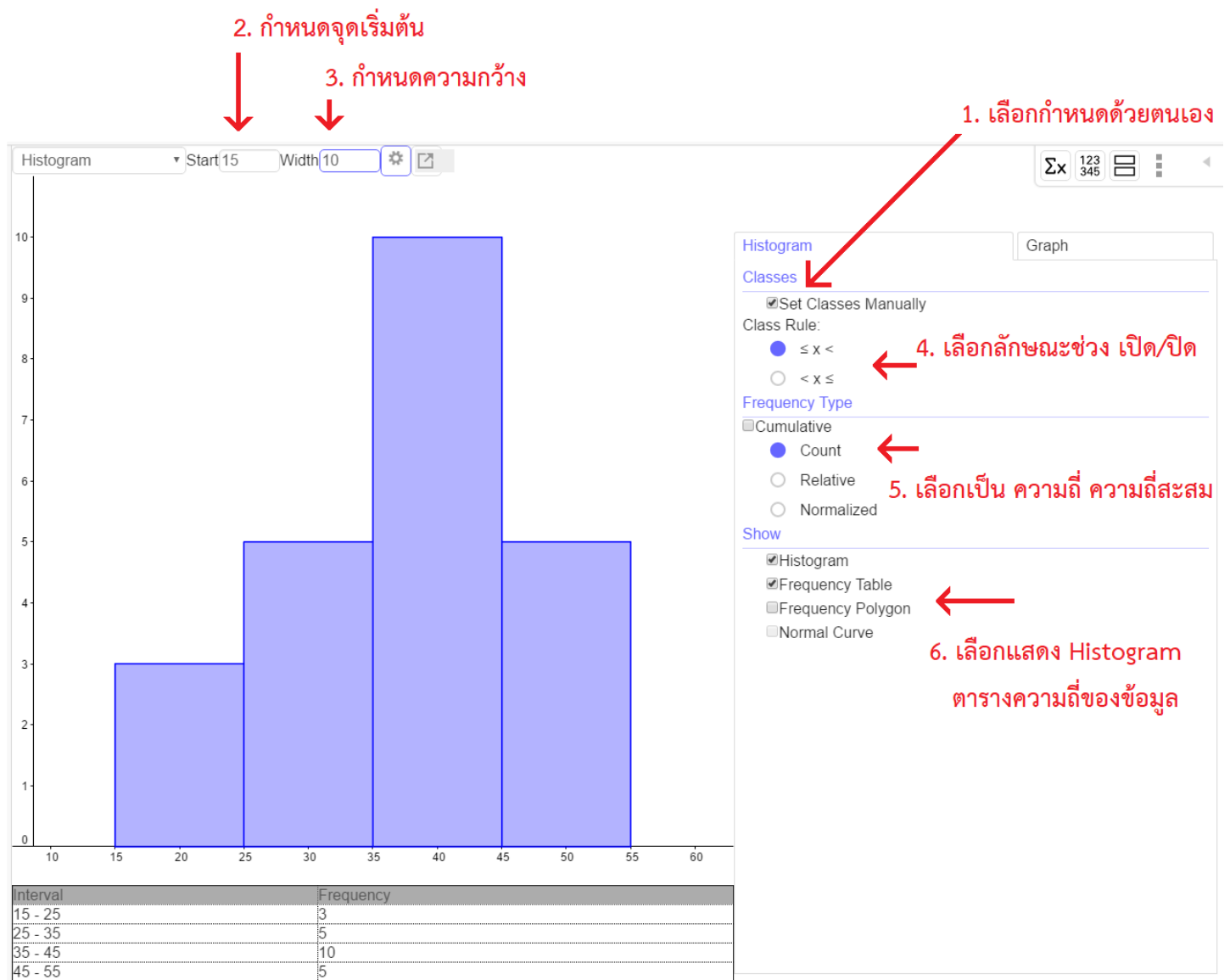
การปรับแต่งกราฟ Histogram โดยใช้เมนู Setting

1. คลิกที่ setting จะได้เมนูย่อยของการปรับแต่ง Histogram

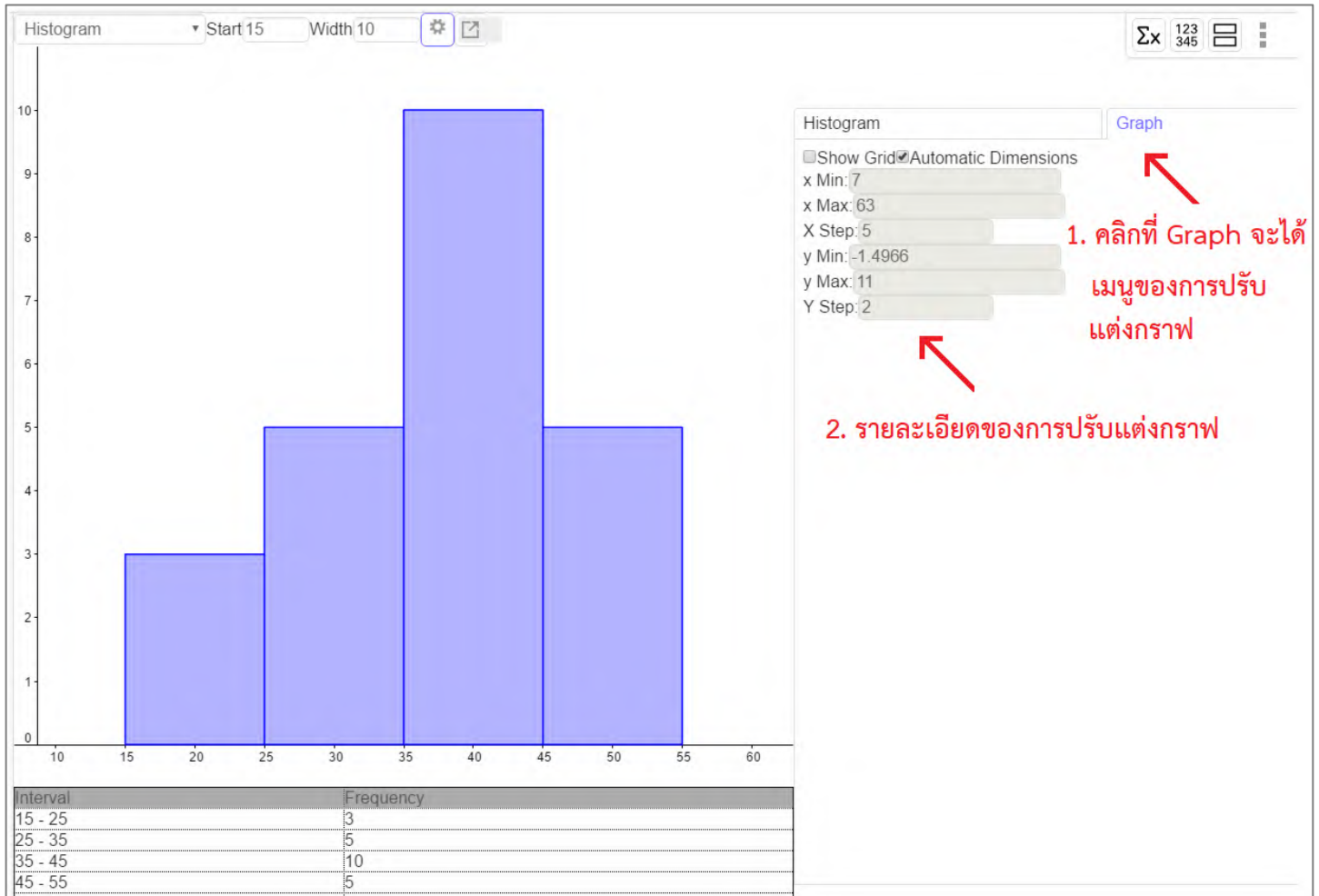
2. เมนูย่อยของการปรับแต่ง Histogram



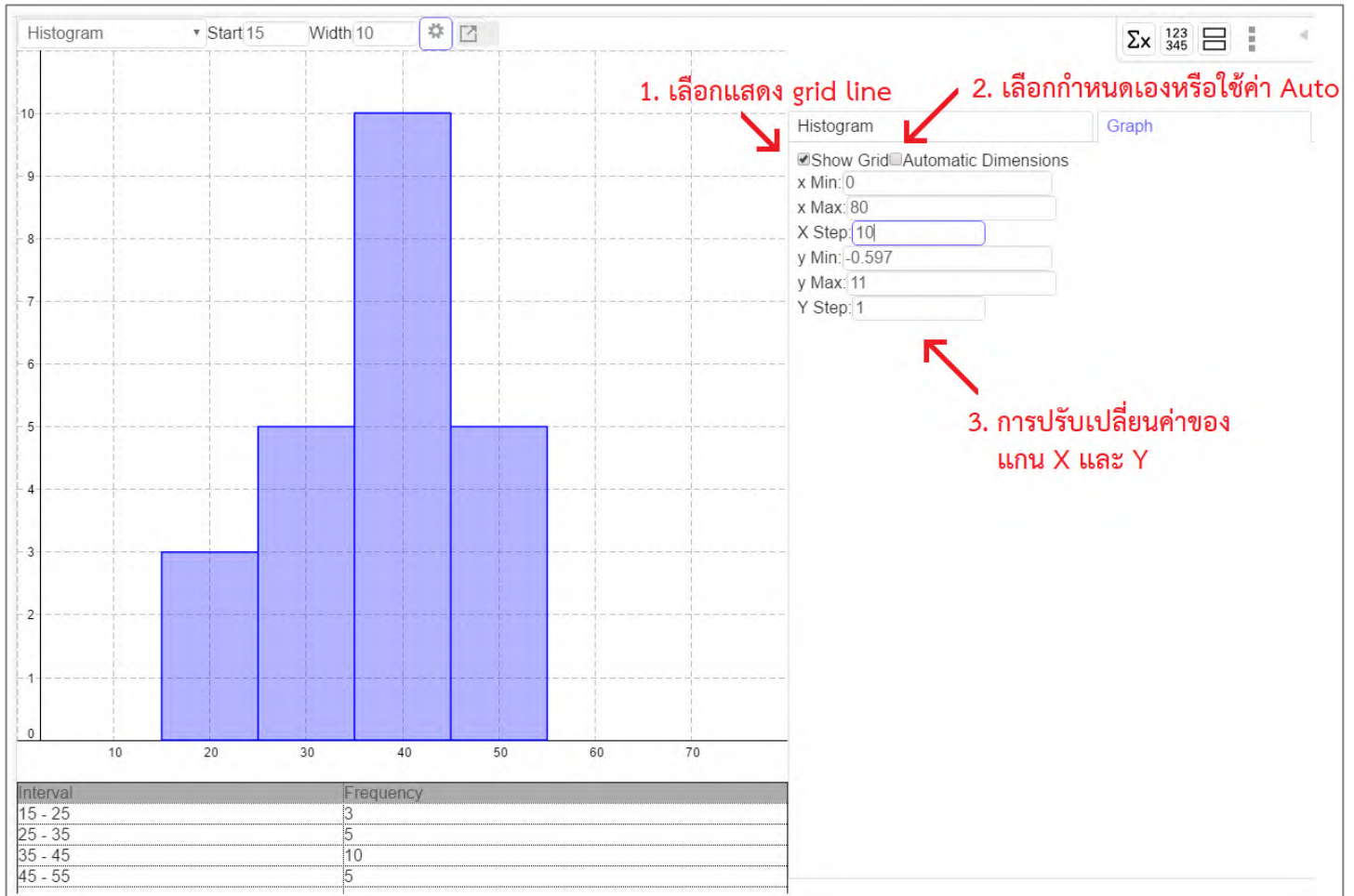
การปรับแต่งกราฟ Histogram



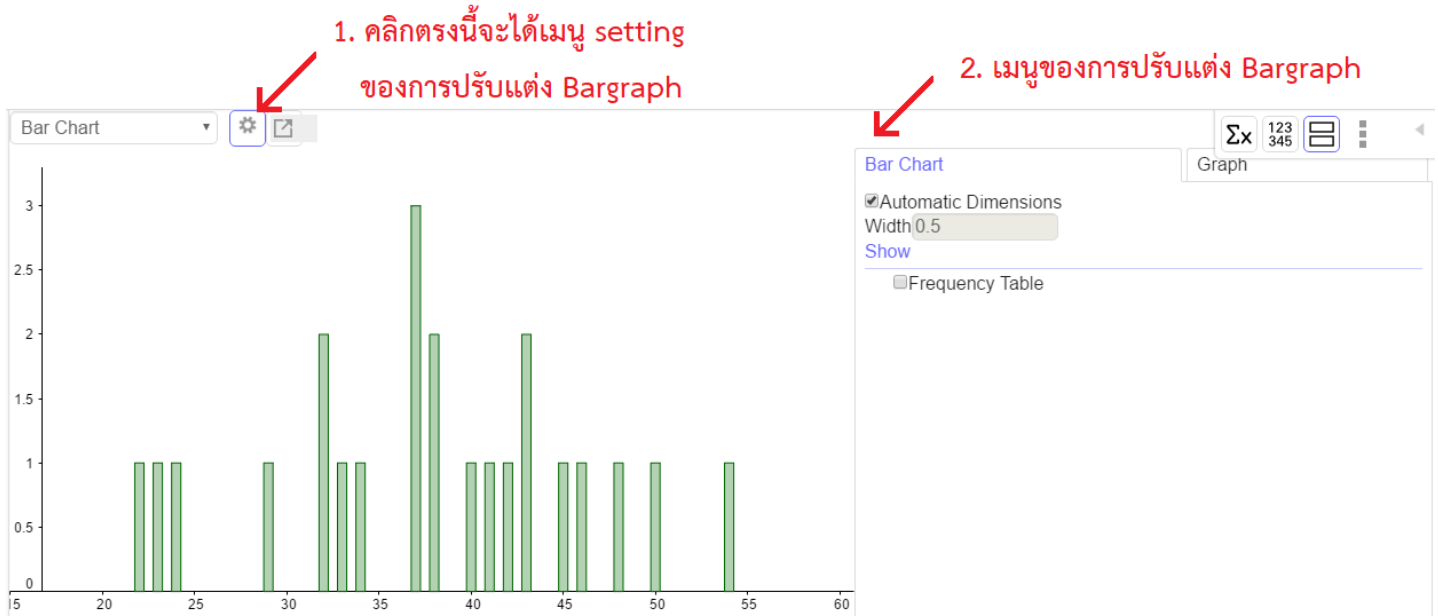
การปรับแต่งกราฟ Histogram โดยใช้เมนู Graph



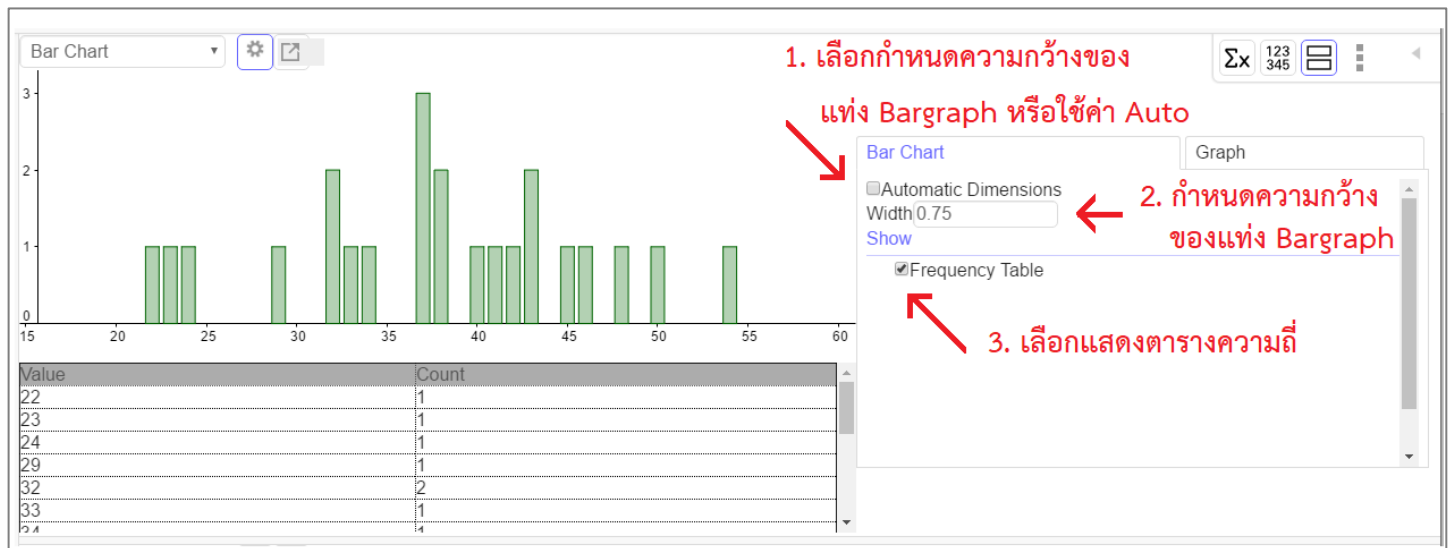
ตัวอย่างเช่น



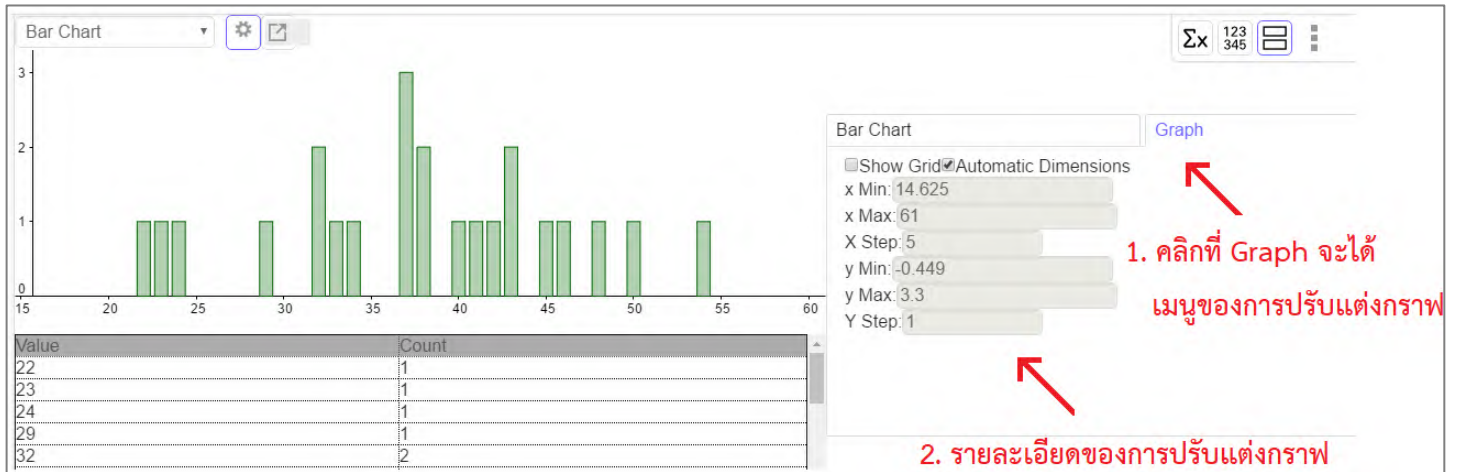
การปรับค่าของการเขียนกราฟแบบ Bargraph



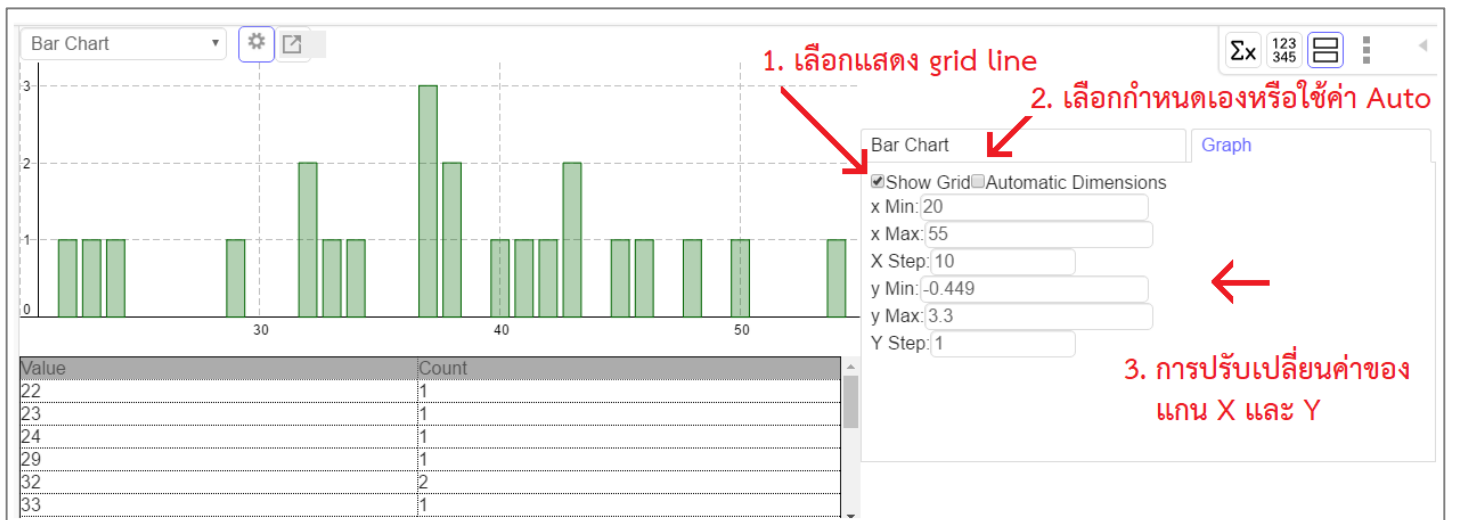
ตัวอย่างเช่น



การปรับแต่งกราฟ Bargraph โดยใช้เมนู Graph



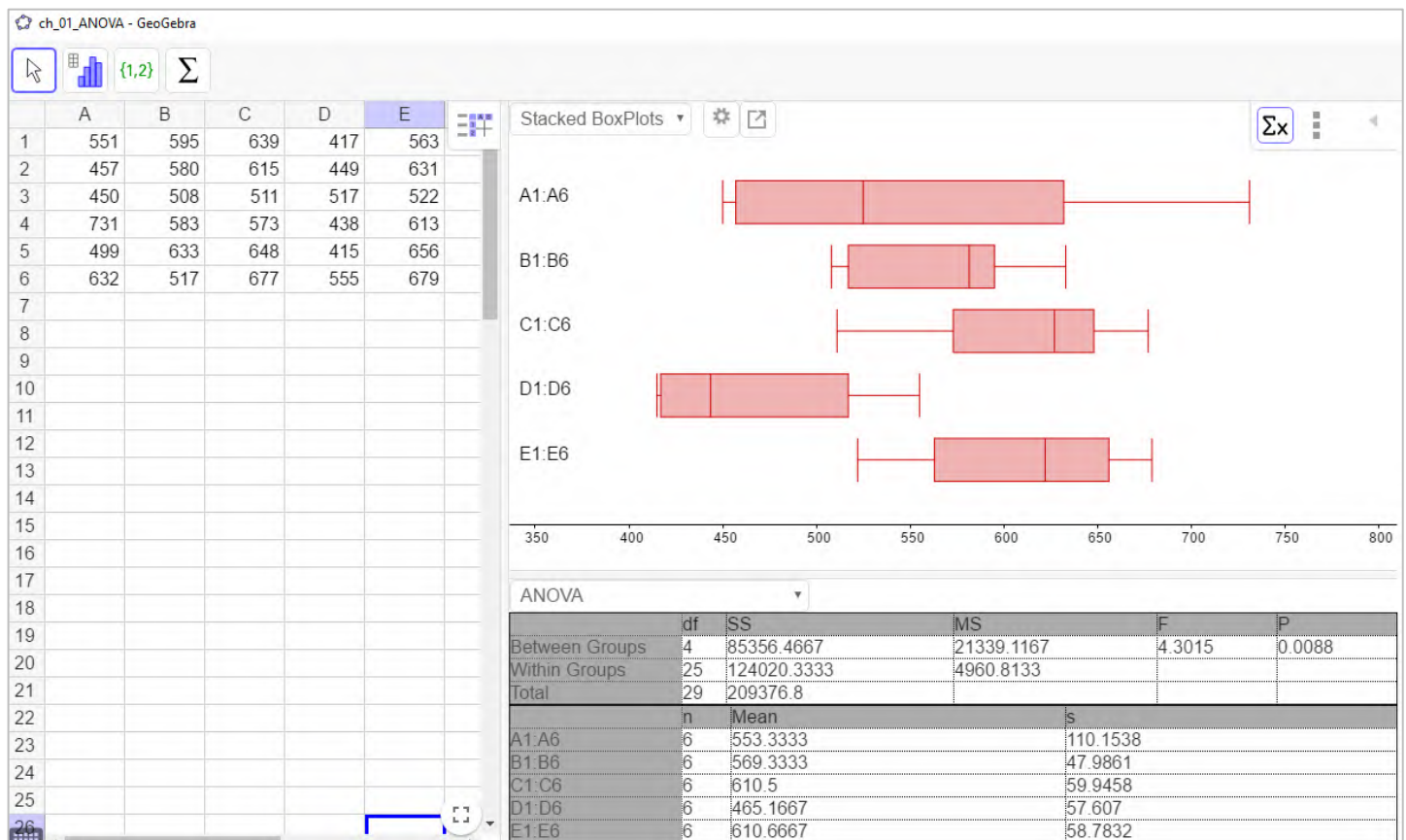
ตัวอย่างเช่น



ตัวอย่างการวิเคราะห์ความแปรปรวนและการสร้างตาราง ANOVA เช่น วิศวกรผู้หนึ่งต้องการทดสอบว่าค่าเฉลี่ยของการดูดความชื้นของคอนกรีตอัดแรง 5 ชนิดเท่ากันหรือไม่ ได้นำตัวอย่างคอนกรีตอัดแรงมาทดลองชนิดละ 6 หน่วย นาน 48 ชั่วโมง บันทึกผลการดูดความชื้นได้ค่าการดูดความชื้น (หน่วย : น้ำหนัก %) ดังนี้

ชนิดที่ 1	ชนิดที่ 2	ชนิดที่ 3	ชนิดที่ 4	ชนิดที่ 5
551	595	639	417	563
457	580	615	449	631
450	508	511	517	522
731	583	573	438	613
499	633	648	415	656
632	517	677	555	679

ผลการวิเคราะห์ความแปรปรวนพร้อมตาราง ANOVA เป็นดังนี้



หมายเหตุ: แฟ้มข้อมูลชื่อ ch_01_ANOVA.ggb

หมายเหตุ ผลการคำนวณของ Excel

L21	:							
	A	B	C	D	E	F	G	
1	Anova: Single Factor							
2								
3	SUMMARY							
4	<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>			
5	type1	6	3320	553.3333333	12133.86667			
6	type2	6	3416	569.3333333	2302.666667			
7	type3	6	3663	610.5	3593.5			
8	type4	6	2791	465.1666667	3318.566667			
9	type5	6	3664	610.6666667	3455.466667			
10								
11								
12	ANOVA							
13	<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>	
14	Between Groups	85356.46667	4	21339.11667	4.301535904	0.008751641	2.75871047	
15	Within Groups	124020.3333	25	4960.813333				
16								
17	Total	209376.8	29					