1.3 ฮิสโทแกรม

การนำเสนอข้อมูลเชิงปริมาณด้วยแผนภาพจุดและแผนภาพต้น–ใบ จะทำให้เห็นข้อมูลทุกตัวที่เก็บรวบรวมได้และลักษณะการกระจายของข้อมูล แต่ในกรณีที่ข้อมูลมีจำนวนมาก ๆ การนำเสนอข้อมูลด้วยรูปแบบดังกล่าวนี้ไม่สะดวกและไม่เป็นที่นิยม การนำเสนอข้อมูลด้วย**ฮิสโทแกรม (histogram)** จึงเป็นอีกรูปแบบหนึ่งที่นิยมใช้ และช่วยให้เห็นลักษณะการกระจายของข้อมูลเช่นกัน

ฮิสโทแกรมมีลักษณะคล้ายแผนภูมิแท่ง แต่ใช้แท่งสี่เหลี่ยมมุมฉากแสดงความถี่หรือความถี่สัมพัทธ์ของข้อมูลเชิงปริมาณในแต่ละช่วง ในขณะที่แผนภูมิแท่งใช้สำหรับข้อมูลเชิงคุณภาพและใช้แท่งสี่เหลี่ยมมุมฉากแสดงปริมาณของข้อมูลซึ่งมีเพียงค่าเดียว

มุมคณิต

ความถี่สัมพัทธ์ (relative frequency) ของค่าที่เป็นไปได้ค่าใดหรือช่วงใด คือ อัตราส่วนระหว่างความถี่ของค่านั้นหรือของช่วงนั้นกับผลรวมของความถี่ทั้งหมด ความถี่สัมพัทธ์อาจแสดงในรูปเศษส่วนหรือทศนิยมหรือร้อยละก็ได้

การสร้างฮิสโทแกรม ทำได้ดังนี้

1. แบ่งข้อมูลออกเป็นช่วง ช่วงละเท่า ๆ กัน ในกรณีของข้อมูลเชิงปริมาณแบบไม่ต่อเนื่องที่มีจำนวนไม่มาก ใช้ข้อมูลแต่ละตัวในการสร้างได้เลย โดยไม่จำเป็นต้องแบ่งเป็นช่วงก็ได้
2. นับจำนวนข้อมูลแต่ละตัวในแต่ละช่วง จำนวนดังกล่าวจะเป็นความถี่ของข้อมูลในช่วงนั้น แล้วสร้างตารางระบุความถี่ของข้อมูลในช่วงนั้น ๆ ซึ่งเรียกว่า ตารางแจกแจงความถี่
3. เขียนแสดงค่าของข้อมูลหรือจุดปลายของช่วงบนแกนนอน แล้วเขียนแท่งสี่เหลี่ยมมุมฉากบนตำแหน่งที่แสดงค่าของข้อมูล โดยให้ความสูงของแท่งเท่ากับความถี่หรือเปอร์เซ็นต์ของความถี่

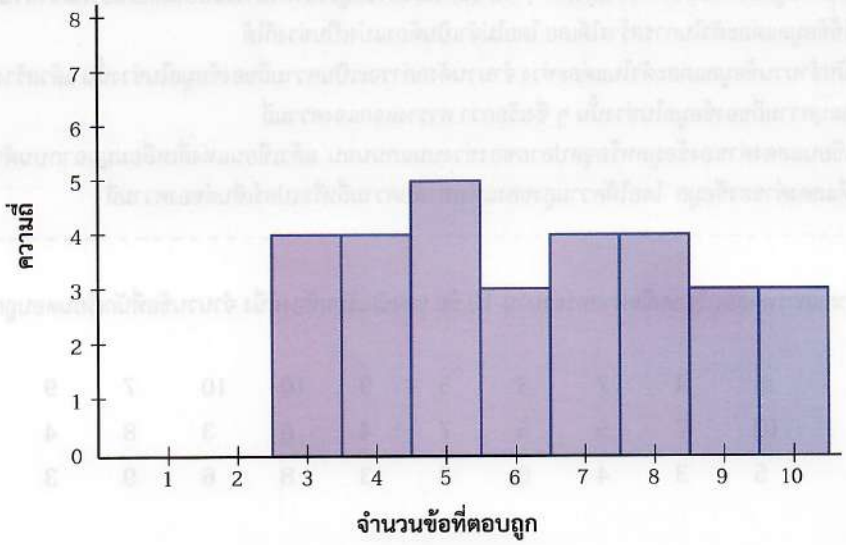
พิจารณาผลการทดสอบวิชาคณิตศาสตร์จำนวน 10 ข้อ ของนักเรียนห้องหนึ่ง จำนวนข้อที่นักเรียนตอบถูก เป็นดังนี้

5	4	7	8	5	9	10	10	7	9
10	7	5	5	7	4	6	3	8	4
5	3	4	6	8	3	8	6	9	3

จากข้อมูล สร้างตารางแจกแจงความถี่ของจำนวนข้อที่นักเรียนตอบถูก ได้ดังนี้

จำนวนข้อที่ตอบถูก	ความถี่
3	4
4	4
5	5
6	3
7	4
8	4
9	3
10	3

เขียนแสดงจำนวนข้อที่นักเรียนตอบถูกบนแกนนอน และเขียนแท่งสี่เหลี่ยมมุมฉากแสดงความถี่ของจำนวนข้อที่นักเรียนตอบถูก จะได้ฮิสโทแกรมดังรูป



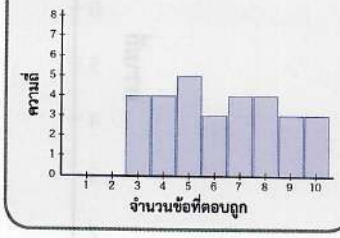
จากฮิสโทแกรม จะเห็นได้ชัดว่า

- นักเรียนห้องนี้ตอบถูก ตั้งแต่ 3 ถึง 10 ข้อ
- นักเรียนส่วนมากตอบถูก 5 ข้อ
- นักเรียนที่ตอบถูกตั้งแต่ 5 ข้อขึ้นไปมี
  $5 + 3 + 4 + 4 + 3 + 3 = 22$ คน
- ลักษณะของข้อมูลโดยรวมมีการกระจายค่อนข้างสม่ำเสมอ



มุมคณิต

ในบางครั้งการเขียนฮิสโทแกรมของข้อมูลเชิงปริมาณที่ไม่ต่อเนื่อง อาจมีช่องว่างระหว่างแท่งได้ เพื่อแสดงให้เห็นลักษณะว่าเป็นฮิสโทแกรมของข้อมูลเชิงปริมาณที่ไม่ต่อเนื่อง

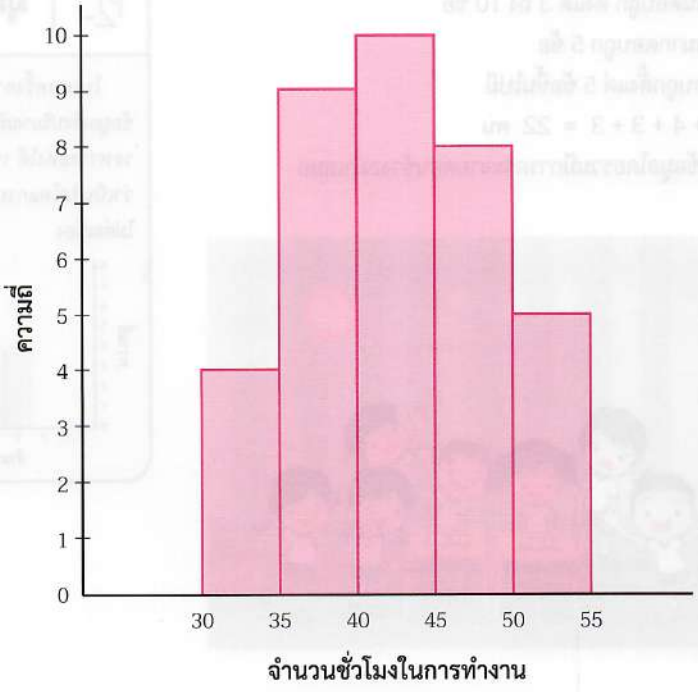


นอกจากนี้เรายังสามารถเขียนฮิสโทแกรมจากข้อมูลที่อยู่ในรูปตารางแจกแจงความถี่ที่เป็นช่วง เช่น จากตารางแสดงความถี่ของจำนวนชั่วโมงในการทำงานแต่ละสัปดาห์ของพนักงาน 36 คน

จำนวนชั่วโมงในการทำงาน	ความถี่
ตั้งแต่ 30 แต่น้อยกว่า 35	4
ตั้งแต่ 35 แต่น้อยกว่า 40	9
ตั้งแต่ 40 แต่น้อยกว่า 45	10
ตั้งแต่ 45 แต่น้อยกว่า 50	8
ตั้งแต่ 50 แต่น้อยกว่า 55	5



สามารถเขียนฮิสโทแกรมแสดงจำนวนชั่วโมงในการทำงานได้ดังนี้



จากฮิสโทแกรม จะเห็นได้ชัดว่า การกระจายของข้อมูลมีลักษณะค่อนข้างสมมาตร ความถี่สูงสุดของข้อมูลอยู่ในช่วงตรงกลาง คือ ตั้งแต่ 40 ชั่วโมงแต่น้อยกว่า 45 ชั่วโมง และความถี่ในช่วงถัดมาจากตรงกลางไปทางซ้าย และจากตรงกลางไปทางขวา ใกล้เคียงกันในแต่ละช่วง

แบบฝึกหัด 1.3

1. จากการสำรวจจำนวนพี่น้องในครอบครัวของนักเรียนแต่ละคนในห้อง จำนวน 32 คน ได้ผลดังนี้

4	1	3	3	2	3	4	3
2	3	2	3	1	1	2	3
1	1	2	4	2	1	2	1
2	2	4	2	4	1	1	3



1) จงสร้างฮิสโทแกรม

2) นักเรียนสรุปข้อมูลได้ว่าอย่างไรจากฮิสโทแกรมที่สร้างในข้อ 1)

2. จากการสำรวจพฤติกรรมการดื่มกาแฟของประชาชนในประเทศต่าง ๆ 30 ประเทศ โดยเก็บรวบรวมข้อมูลน้ำหนักของกาแฟ โดยเฉลี่ยที่แต่ละคนบริโภคในเวลา 1 ปี ได้ผลดังนี้

ประเทศ	น้ำหนักของกาแฟโดยเฉลี่ยที่แต่ละคนบริโภค (กิโลกรัมต่อปี)	ประเทศ	น้ำหนักของกาแฟโดยเฉลี่ยที่แต่ละคนบริโภค (กิโลกรัมต่อปี)	ประเทศ	น้ำหนักของกาแฟโดยเฉลี่ยที่แต่ละคนบริโภค (กิโลกรัมต่อปี)
ฟินแลนด์	9.6	แคนาดา	3.4	สวิตเซอร์แลนด์	3.9
เอสโตเนีย	4.2	เกาหลีใต้	2.6	ออสเตรีย	5.5
โปรตุเกส	2.6	บัลแกเรีย	2.3	กรีซ	2.4
สโลวีเนีย	6.1	นอร์เวย์	7.2	สหรัฐอเมริกา	3.1
ไทย	1.6	อิสราเอล	1.8	สิงคโปร์	2.1
ฝรั่งเศส	3.2	เยอรมนี	5.2	เดนมาร์ก	5.3
สเปน	3	คอสตาริกา	3.7	โปแลนด์	3.1
เซอร์เบีย	5.4	ออสเตรเลีย	2.6	โรมาเนีย	1.9
อิตาลี	3.4	กัวเตมาลา	2	เนเธอร์แลนด์	6.7
บราซิล	4.8	ฮังการี	3.1	เบลเยียม	4.9

ที่มา : http://www.caffeineinformer.com สืบค้นเมื่อ 8 เมษายน 2561

1) จงสร้างตารางแจกแจงความถี่ เมื่อกำหนดช่วงให้ดังนี้

ตั้งแต่ 0.7 แต่น้อยกว่า 2.2
ตั้งแต่ 2.2 แต่น้อยกว่า 3.7
ตั้งแต่ 3.7 แต่น้อยกว่า 5.2
ตั้งแต่ 5.2 แต่น้อยกว่า 6.7
ตั้งแต่ 6.7 แต่น้อยกว่า 8.2
ตั้งแต่ 8.2 แต่น้อยกว่า 9.7



2) จงสร้างฮิสโทแกรมจากตารางแจกแจงความถี่ที่สร้างได้ในข้อ 1)

3) นักเรียนสรุปข้อมูลจากฮิสโทแกรมได้ว่าอย่างไร

3. จากการสำรวจของ General Social Survey ในปี ค.ศ. 2016 เกี่ยวกับจำนวนชั่วโมงเฉลี่ยในแต่ละวันที่ใช้ในการดูทีวีของชาวอเมริกัน โดยมีผู้ตอบแบบสำรวจทั้งหมด 1,883 คน ได้ผลดังนี้

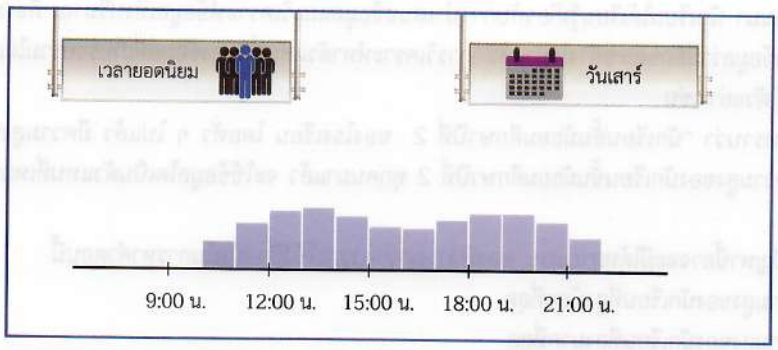
เวลาที่ใช้ในการดูทีวี (ชั่วโมงต่อวัน)	จำนวนคน
0	162
1	370
2	486
3	297
4	221
5	113
6	92
7	22
8	46
9	6
10	23
11	1
12	25

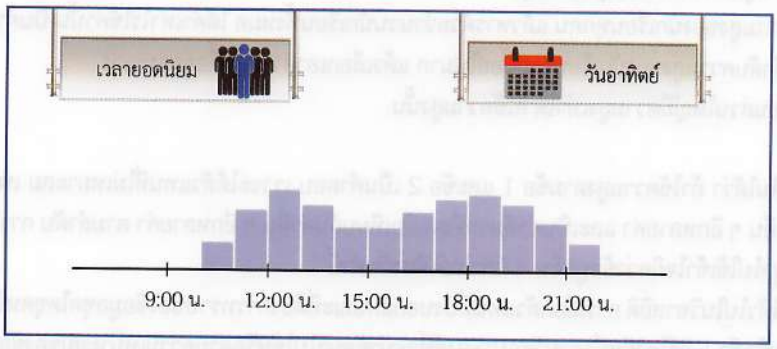
เวลาที่ใช้ในการดูทีวี (ชั่วโมงต่อวัน)	จำนวนคน
13	1
14	1
15	4
16	2
17	0
18	1
19	0
20	5
21	0
22	0
23	0
24	5

1) จงสร้างฮิสโทแกรมแสดงจำนวนชั่วโมงที่ใช้ในการดูทีวี โดยให้แกนตั้งแสดงเปอร์เซ็นต์ของจำนวนคน
2) ข้อมูลจากหนังสือพิมพ์ The Telegraph ประจำวันที่ 10 กรกฎาคม ค.ศ. 2012 อ้างว่า นักวิทยาศาสตร์กล่าวว่า "การจำกัดเวลาในการดูทีวีไม่เกินสองชั่วโมงต่อวัน สามารถยืดชีวิตให้ยืนยาวได้อีกหนึ่งปีครึ่ง" หากข้อมูลนี้ยังเชื่อถือได้ กลุ่มของผู้ที่ตอบแบบสำรวจที่มีโอกาสมีชีวิตยืนยาวได้อีกหนึ่งปีครึ่ง จากการจำกัดเวลาในการดูทีวีไม่เกินสองชั่วโมงต่อวันคิดเป็นร้อยละเท่าใดของผู้ที่ตอบแบบสำรวจ
3) นักเรียนจะสรุปเกี่ยวกับผลการสำรวจนี้ได้ว่าอย่างไร



4. ออโอ้ค้นหาข้อมูลของร้านอาหารที่ชื่นชอบแห่งหนึ่ง ได้ข้อมูลดังนี้





1) นักเรียนคิดว่า เวลาเปิดให้บริการของร้านแห่งนี้คือช่วงเวลาใด
2) หากต้องการจะไปรับประทานอาหารที่ร้านแห่งนี้ในวันเสาร์หรือวันอาทิตย์ นักเรียนจะไปในช่วงเวลาใด เพราะเหตุใด
3) จำนวนลูกค้าที่เข้ามารับประทานอาหารที่ร้านอาหารแห่งนี้ในวันเสาร์และวันอาทิตย์แตกต่างกันหรือไม่ อย่างไร

