

แบบฝึกหัด หน่วยที่ 2 เรื่อง การเก็บรวบรวมและสำรวจข้อมูล (ต่อ)

1. ให้ทำเครื่องหมาย ✓ หน้าข้อที่กล่าวถูกต้อง

- ☐ ตารางที่มีรูปแบบ wide table เหมาะสำหรับการแปลงข้อมูลด้วยการรวมข้อมูล
- ☐ เครื่องที่ใช้ในการรวมข้อมูลคือ pivotable
- ☐ ในการแปลงข้อมูลต้องทำครบทั้ง 3 วิธี คือ การเพิ่มข้อมูล การลดข้อมูล และการรวมข้อมูล
- ☐ การเชื่อมโยงข้อมูลเป็นการนำข้อมูลจากแหล่งเดียวกันมาจัดกลุ่มตามเงื่อนไขที่กำหนด
- ☐ การสำรวจข้อมูล คือ เป็นการทดลองวาดแผนภาพ หรือกราฟข้อมูลในรูปแบบต่าง ๆ เพื่อพิจารณาภาพรวมของข้อมูล
- ☐ ข้อมูลส่วนบุคคล หมายถึง ข้อมูลเกี่ยวกับสิ่งเฉพาะตัวของบุคคล เช่น การศึกษา ฐานะการเงิน

2. เครื่องมือใดที่เหมาะสมสำหรับใช้สำรวจข้อมูลที่ลักษณะต่อไปนี้

แสดงข้อมูลที่ประกอบไปด้วยกลุ่มย่อยหลาย ๆ กลุ่ม

การเชื่อมโยงช่วยพยากรณ์แนวโน้มในอนาคต

สำรวจความสัมพันธ์ระหว่างสองแอตทริบิวต์

แสดงความถี่ของสิ่งที่สนใจ

3. จากข้อมูลในตารางควรรนำเสนอแผนภาพแบบใด

ลำดับที่	ยอดขายรองเท้า	จำนวน (คู่)
1	เดือน มกราคม-มีนาคม	258
2	เดือน เมษายน - มิถุนายน	485
3	เดือน กรกฎาคม - กันยายน	153
4	เดือนตุลาคม - ธันวาคม	254

.....

ลำดับที่	อุณหภูมิ	ยอดขาดไอสกรีม
1	21	524
2	28	600
3	30	650

.....

รายการ	2014	2015	2016	2017	2018	2019	2020	2021	2022
การก่อสร้างใหม่	213,933	196,334	218,986	355,698	415,320	456,00	487,145	497,214	501,447
ที่มีอยู่เดิม	123,345	145,899	189,000	200,433	340,210	345,500	347,785	358,951	368,411

.....

ฝนตกรายวัน	ฝุ่นละออง
4.1	122
4.3	117
5.7	112
5.4	114
5.9	110
5.0	114
3.6	128
1.9	137
7.3	104

.....

เดือน	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.	มิ.ย.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.
Smartphone	10	20	11	14	58	14	12	22	14	17	18
Tablet	12	15	17	13	21	23	14	16	17	18	11
Notebook	20	23	27	28	21	25	24	23	25	26	27
Desktop	5	6	7	8	4	2	1	5	6	4	5
Computer Set	8	7	2	5	4	8	6	4	7	8	9
Monitor	12	45	85	14	75	26	14	18	18	17	15
Mouse	20	30	25	25	24	14	72	45	46	35	26
Keyboard	25	26	24	28	35	24	12	25	27	28	31

.....

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

.....

4. ใช้ข้อมูลต่อไปนี้ตอบคำถาม

มีฐานข้อมูลชื่อ weight_height.csv ซึ่งมีข้อมูลดังนี้

No	Name_surname	weight	height
1	Malee Rukma	65	178
:	:	:	:
50	Rose Mala	70	180

ถ้าต้องนำฐานข้อมูลดังกล่าวในตารางมาสำรวจข้อมูลโดยการเขียนโปรแกรมเป็นแผนภาพการกระจายเพื่อดูความสัมพันธ์ระหว่างน้ำหนักและส่วนสูงโดยให้แกน x มีชื่อว่า Weight และ แกน y มีชื่อว่า Height

```

import pandas as pd
import matplotlib.pyplot as plt
df= pd.read_csv('_____')
x=df. _____
y=df. _____
plt.scatter(x,y)
plt.xlabel('_____')
plt.ylabel('_____');

```

5. ในการสำรวจข้อมูลโดยการเขียนโปรแกรมด้วยภาษาไพทอน

ไลบรารีใดที่ใช้ในการจัดการฐานข้อมูล

ไลบรารีใดที่ใช้ในการเรียกใช้การสร้างแผนภาพ