

1. ก.) เชิงคุณภาพ ชนิด nominal แบบ binary
 ข.) เชิงปริมาณ interval แบบ continuous
 ค.) เชิงคุณภาพ ordinal แบบ discrete
 ง.) เชิงปริมาณ ratio แบบ continuous
 จ.) เชิงคุณภาพ ordinal แบบ discrete
 ฉ.) เชิงปริมาณ interval แบบ continuous
 ช.) เชิงปริมาณ interval แบบ discrete
 ซ.) เชิงคุณภาพ nominal แบบ discrete
 ฌ.) เชิงปริมาณ ratio แบบ continuous
 ญ.) เชิงปริมาณ interval แบบ continuous
 ฎ.) เชิงคุณภาพ ordinal แบบ discrete
 ฏ.) เชิงปริมาณ ratio แบบ continuous

2. - ต้องการรู้ช่วงอายุ
 - ต้องการรู้ช่วงปีจบ
 - ต้องการรู้วิชาที่น่านจะเรียนในปัจจุบัน

3. daily temperature เพราะ ข้อมูลจากพื้นที่หนึ่งที่ได้นั้นจะมีค่าใกล้เคียงกลับพื้นที่ ๆ อยู่ใกล้ ๆ

4. ก. - มีประสิทธิภาพต่ำจากการที่ต้องนำตัวเข้ามาประมวลผลร่วมด้วย
 - ได้เพื่อนบ้านที่เป็นตัวซ้ำกับตัวต้น
 - หากมีตัวใกล้ ๆ เป็นตัวซ้ำอาจทำให้เพื่อนบ้านที่ได้มีแต่นั้น ๆ
 ข. จัดรวมกลุ่มตัวซ้ำเป็นกลุ่มเดียวกันแล้วให้บันทึกกลุ่มนั้นเป็น data object ตัวหนึ่ง

5.)

ก. cosine

$$\langle x, y \rangle = 1 \cdot 2 + 1 \cdot 2 + 1 \cdot 2 + 1 \cdot 2 = 8$$

$$\|x\| = (1^2 + 1^2 + 1^2 + 1^2)^{1/2} = 2 \quad \|y\| = (2^2 + 2^2 + 2^2 + 2^2)^{1/2} = 4$$

$$\text{Cosine} = 8 / (2 \cdot 4) = 1$$

$$\text{Correlation mean}(x) = (1+1+1+1)/4 = 1 \quad \text{mean}(y) = (2+2+2+2)/4 = 2$$

$$S_x = (1/4 - 1 \cdot ((1-1)^2 + (1-1)^2 + (1-1)^2 + (1-1)^2))^{1/2} = 0 \quad S_y = (1/4 - 1 \cdot ((2-2)^2 + (2-2)^2 + (2-2)^2 + (2-2)^2))^{1/2} = 0$$

$$S_{xy} = (1/4 - 1 \cdot ((1-1)(2-2) + (1-1)(2-2) + (1-1)(2-2) + (1-1)(2-2)))^{1/2} = 0$$

$$\text{Correlation} = 0/0*0 = 0$$

Jaccard

$$f_{1,2} = 4$$

$$J = 0/4 = 0$$

Euclidean distance

$$d(x,y) = ((1-2)^2 + (1-2)^2 + (1-2)^2 + (1-2)^2)^{1/2} = 2$$

u.) cosine

$$\langle x,y \rangle = 0*1 + 1*0 + 0*1 + 1*0 = 0$$

$$\|x\| = (0^2 + 1^2 + 0^2 + 1^2)^{1/2} = 2^{1/2} \quad \|y\| = (1^2 + 0^2 + 1^2 + 0^2)^{1/2} = 2^{1/2}$$

$$\text{Cosine} = 0/2^{1/2}*2^{1/2} = 0$$

$$\text{Correlation mean}(x) = (1+0+1+0)/4 = 1/2 \quad \text{mean}(y) = (0+1+0+1)/4 = 1/2$$

$$S_x = (1/4 - 1*((1-1/2)^2 + (0-1/2)^2 + (1-1/2)^2 + (0-1/2)^2))^{1/2} = 1/3^{1/2} \quad S_y = (1/4 - 1*((0-1/2)^2 + (1-1/2)^2 + (0-1/2)^2 + (1-1/2)^2))^{1/2} = 1/3^{1/2}$$

$$S_{xy} = (1/4 - 1*((1-1/2)(0-1/2) + (0-1/2)(1-1/2) + (1-1/2)(0-1/2) + (0-1/2)(1-1/2)))^{1/2} = (-1/3)^{1/2}$$

$$\text{Correlation} = (-1/3)^{1/2} / (1/3)^{1/2} * (1/3)^{1/2} = 1/(-1/3)^{1/2}$$

Jaccard

$$f_{1,0} = 2 \quad f_{0,1} = 2$$

$$J = 0/2+2 = 0$$

Euclidean distance

$$d(x,y) = ((1-0)^2 + (0-1)^2 + (1-0)^2 + (0-1)^2)^{1/2} = 2$$

u. cosine

$$\langle x,y \rangle = 0*1 + -1*0 + 0*-1 + 1*0 = 0$$

$$\|x\| = (0^2 + 1^2 + 0^2 + (-1)^2)^{1/2} = 2^{1/2} \quad \|y\| = (1^2 + 0^2 + (-1)^2 + 0^2)^{1/2} = 2^{1/2}$$

$$\text{Cosine} = 0/2^{1/2}*2^{1/2} = 0$$

$$\text{Correlation mean}(x) = (0+1+0-1)/4 = 0 \quad \text{mean}(y) = (1+0-1+0)/4 = 0$$

$$S_x = (1/4 - 1*((0-0)^2 + (1-0)^2 + (0-0)^2 + (-1-0)^2))^{1/2} = 2/3^{1/2} \quad S_y = (1/4 - 1*((1-0)^2 + (0-0)^2 + (-1-0)^2 + (0-0)^2))^{1/2} = 1/3^{1/2}$$

$$S_{xy} = (1/4 - 1*((0-0)(1-0) + (1-0)(0-0) + (0-0)(-1-0) + (1-0)(0-0)))^{1/2} = 0$$

$$\text{Correlation} = 0/1/3^{1/2}*1/3^{1/2} = 0$$

Jaccard

$$F_{0,1} = 1 \quad f_{1,0} = 1 \quad f_{0,-1} = 1 \quad f_{-1,0} = 1$$

$$J = 0/(1+1+1+1) = 0$$

Euclidean distance

$$d(x,y) = ((0-1)^2 + (1-0)^2 + (0-(-1))^2 + (-1-0)^2)^{1/2} = 2$$

6. n.

| x_j | $P(X=x_j)$ | $-P(X=x_j)\log_2 P(X=x_j)$ |
|--------|------------|----------------------------|
| -7 | 1/6 | 0.43082 |
| -2 | 1/6 | 0.43082 |
| 0 | 1/6 | 0.43082 |
| 1 | 2/6 | 0.52832 |
| 2 | 1/6 | 0.43082 |
| $H(x)$ | | 2.2516 |

| Y_k | $P(Y=y_k)$ | $-P(Y=y_k)\log_2 P(Y=y_k)$ |
|--------|------------|----------------------------|
| 0 | 1/6 | 0.43082 |
| 1 | 2/6 | 0.52832 |
| 4 | 2/6 | 0.52832 |
| 9 | 1/6 | 0.43082 |
| $H(y)$ | | 1.91828 |

| x_j | Y_k | $P(X=x_j, Y=y_k)$ | $-P(X=x_j, Y=y_k) \log_2 P(X=x_j, Y=y_k)$ |
|----------|-------|-------------------|---|
| -7 | 9 | 1/6 | 0.43082 |
| -2 | 4 | 1/6 | 0.43082 |
| 1 | 1 | 1/6 | 0.43082 |
| 0 | 0 | 1/6 | 0.43082 |
| 1 | 4 | 1/6 | 0.43082 |
| 2 | 1 | 1/6 | 0.43082 |
| $H(x,y)$ | | | 2.58492 |

$$I(x,y) = 2.2516 + 1.91828 - 2.58492 = 1.58496$$

7.

| x_j | $P(X=x_j)$ | $-P(X=x_j)\log_2 P(X=x_j)$ |
|--------|------------|----------------------------|
| 1 | 4/4 | 0 |
| $H(x)$ | | 0 |

| Y_k | $P(Y=y_k)$ | $-P(Y=y_k)\log_2 P(Y=y_k)$ |
|--------|------------|----------------------------|
| 2 | 4/4 | 0 |
| $H(y)$ | | 0 |

| x_j | Y_k | $P(X=x_j, Y=y_k)$ | $-P(X=x_j, Y=y_k) \log_2 P(X=x_j, Y=y_k)$ |
|----------|-------|-------------------|---|
| 1 | 2 | 4/4 | 0 |
| $H(x,y)$ | | | 0 |

$$I(x,y) = 0 + 0 - 0 = 0$$