

1.

(a) 水系;國家的資料或田野調查;水系在什麼情況下容易有分支

(b)使用者在不同地點的打卡資訊;可以透過爬蟲或是社群網站取得;要怎麼預測使用者去各個地點的機率

(c)族譜;可以透過戶政事務所的資料得到;若有遺傳性疾病，會如何傳播。

(d)台南市的道路網;從地圖資訊取得;哪些路段容易塞車

(e)學生與課程;可以透過問學或學校的資料來蒐集;什麼樣的學生會選同樣的課程

(f)流行病的傳染;透過健保單位收集資料;若有一個疾病發生，可能會在哪些區域較嚴重。

2.(a)

[0, 0, 0, 0, 1],

[1, 0, 1, 1, 0],

[0, 0, 0, 0, 0],

[0, 1, 1, 0, 1],

[0, 0, 0, 1, 0]

(b)

1: {5},

2: {1,3,4}

3: {}

4: {2, 3, 5}

5: {4}

(c)

top :

[0 1 1 0 0]

[1 0 1 0 0]

[1 1 0 1 0]

[0 0 1 0 1]

[0 0 0 1 0]

bottom :

[0 0 1 1 1 1]

[0 0 0 0 1 0]

[1 0 0 0 1 0]

[1 0 0 0 0 1]

[1 1 1 0 0 0]

[1 0 0 1 0 0]

(d)
$$\frac{2}{\sqrt{4*5}} = 0.4472$$