COMP9318 Assignment

**Student Name: Jiawen Kang**

**ZID: z5055459**

Question 1

1)

|  |  |  |  |
| --- | --- | --- | --- |
| **Location** | **Time** | **Item** | **SUM(Quantity)** |
| **Sydney** | **2005** | **PS2** | **1400** |
| **Sydney** | **2005** | **ALL** | **1400** |
| **Sydney** | **2006** | **PS2** | **1500** |
| **Sydney** | **2006** | **Wii** | **500** |
| **Sydney** | **2006** | **ALL** | **2000** |
| **Sydney** | **ALL** | **PS2** | **2900** |
| **Sydney** | **ALL** | **Wii** | **500** |
| **Sydney** | **ALL** | **ALL** | **3400** |
| **Melbourne** | **2005** | **XBox360** | **1700** |
| **Melbourne** | **2005** | **ALL** | **1700** |
| **Melbourne** | **ALL** | **XBox360** | **1700** |
| **Melbourne** | **ALL** | **ALL** | **1700** |
| **ALL** | **2005** | **PS2** | **1400** |
| **ALL** | **2005** | **XBox360** | **1700** |
| **ALL** | **2005** | **ALL** | **3100** |
| **ALL** | **2006** | **PS2** | **1500** |
| **ALL** | **2006** | **Wii** | **500** |
| **ALL** | **2006** | **ALL** | **2000** |
| **ALL** | **ALL** | **PS2** | **2900** |
| **ALL** | **ALL** | **Wii** | **500** |
| **ALL** | **ALL** | **XBox360** | **1700** |
| **ALL** | **ALL** | **ALL** | **5100** |

2.)

(SELECT Location, Year, Item, Quantity

FROM Sales)

UNION

(SELECT Location, Year, 'ALL', SUM(Quantity)

FROM Sales

GROUP BY(Location, Year))

UNION

(SELECT Location, 'ALL', Item, SUM(Quantity)

FROM Sales

GROUP BY(Location, Item))

UNION

(SELECT 'ALL', Year, Item, SUM(Quantity)

FROM Sales

GROUP BY(Year, Year))

|  |  |  |  |
| --- | --- | --- | --- |
| **Location** | **Time** | **Item** | **SUM(Quantity)** |
| **Sydney** | **2005** | **PS2** | **1400** |
| **Sydney** | **2005** | **ALL** | **1400** |
| **Sydney** | **2006** | **PS2** | **1500** |
| **Sydney** | **2006** | **Wii** | **500** |
| **Sydney** | **2006** | **ALL** | **2000** |
| **Sydney** | **ALL** | **PS2** | **2900** |
| **Sydney** | **ALL** | **Wii** | **500** |
| **Melbourne** | **2005** | **XBox360** | **1700** |
| **Melbourne** | **2005** | **ALL** | **1700** |
| **Melbourne** | **ALL** | **XBox360** | **1700** |
| **ALL** | **2005** | **PS2** | **1400** |
| **ALL** | **2005** | **XBox360** | **1700** |
| **ALL** | **2006** | **PS2** | **1500** |
| **ALL** | **2006** | **Wii** | **500** |

UNION

(SELECT Location, 'ALL', 'ALL', SUM(Quantity)

FROM Sales

GROUP BY(Location))

UNION

(SELECT 'ALL', Year, 'ALL', SUM(Quantity)

FROM Sales

GROUP BY(Year))

UNION

(SELECT 'ALL', 'ALL', Item, SUM(Quantity)

FROM Sales

GROUP BY(Item))

UNION

(SELECT 'ALL', 'ALL', 'ALL', SUM(Quantity)

FROM Sales)

3)

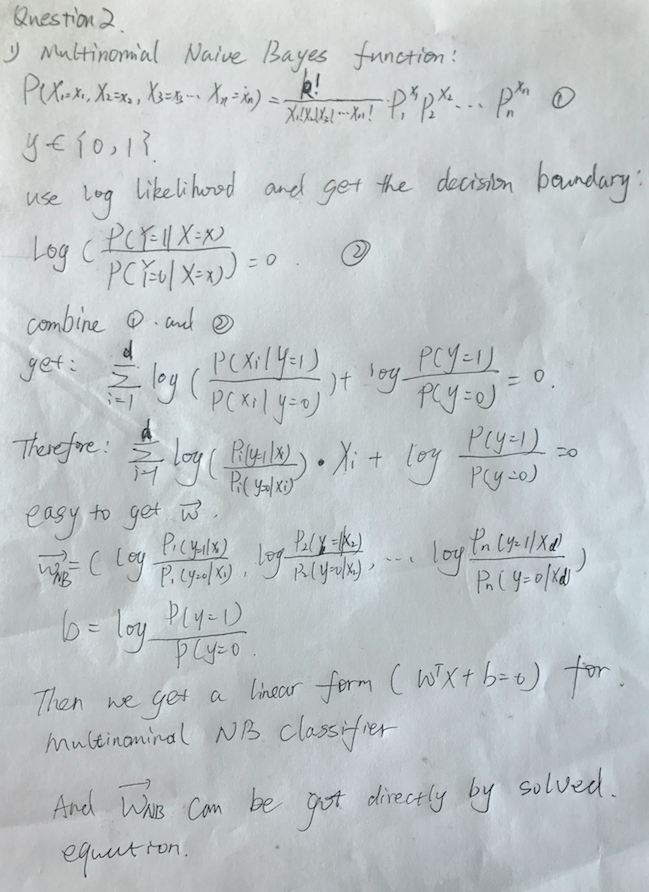
4)

|  |  |
| --- | --- |
| **SUM(Quantity)** | **offset** |
| **1400** | **17** |
| **1400** | **16** |
| **1500** | **21** |
| **500** | **23** |
| **2000** | **20** |
| **2900** | **13** |
| **500** | **15** |
| **3400** | **12** |
| **1700** | **30** |
| **1700** | **28** |
| **1700** | **26** |
| **1700** | **24** |
| **1400** | **5** |
| **1700** | **6** |
| **3100** | **4** |
| **1500** | **9** |
| **500** | **11** |
| **2000** | **8** |
| **2900** | **1** |
| **500** | **3** |
| **1700** | **2** |
| **5100** | **0** |

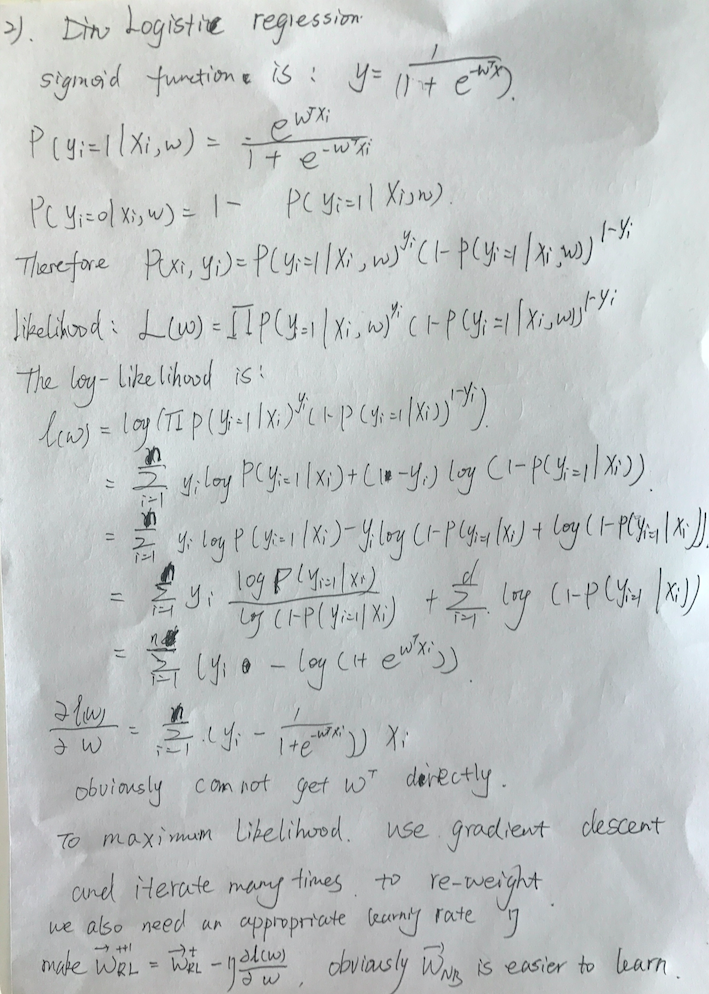
f(Location, Time, Item) = 12\*Location + 4\*Time + Item

for process, in next page

Q2

1

q2

2

Question 3

1)

Data: D is a dataset of n d-dimensional points; k is the number of clusters.

1 Initialize k centers C = [c1, c2, . . . , ck];  
2 canStop ← false;  
3 while canStop = false do

4 Initialize k empty clusters G = [g1, g2, . . . , gk];

5 for each data point p ∈ D do

6 cx ← NearestCenter(p, C);

7 gcx .append(p);

**insert\_1:**

**oldCentres = C;**

**C = [];**

8 for each group g ∈ G do  
9 ci ← ComputeCenter(g);

**insert\_2:**

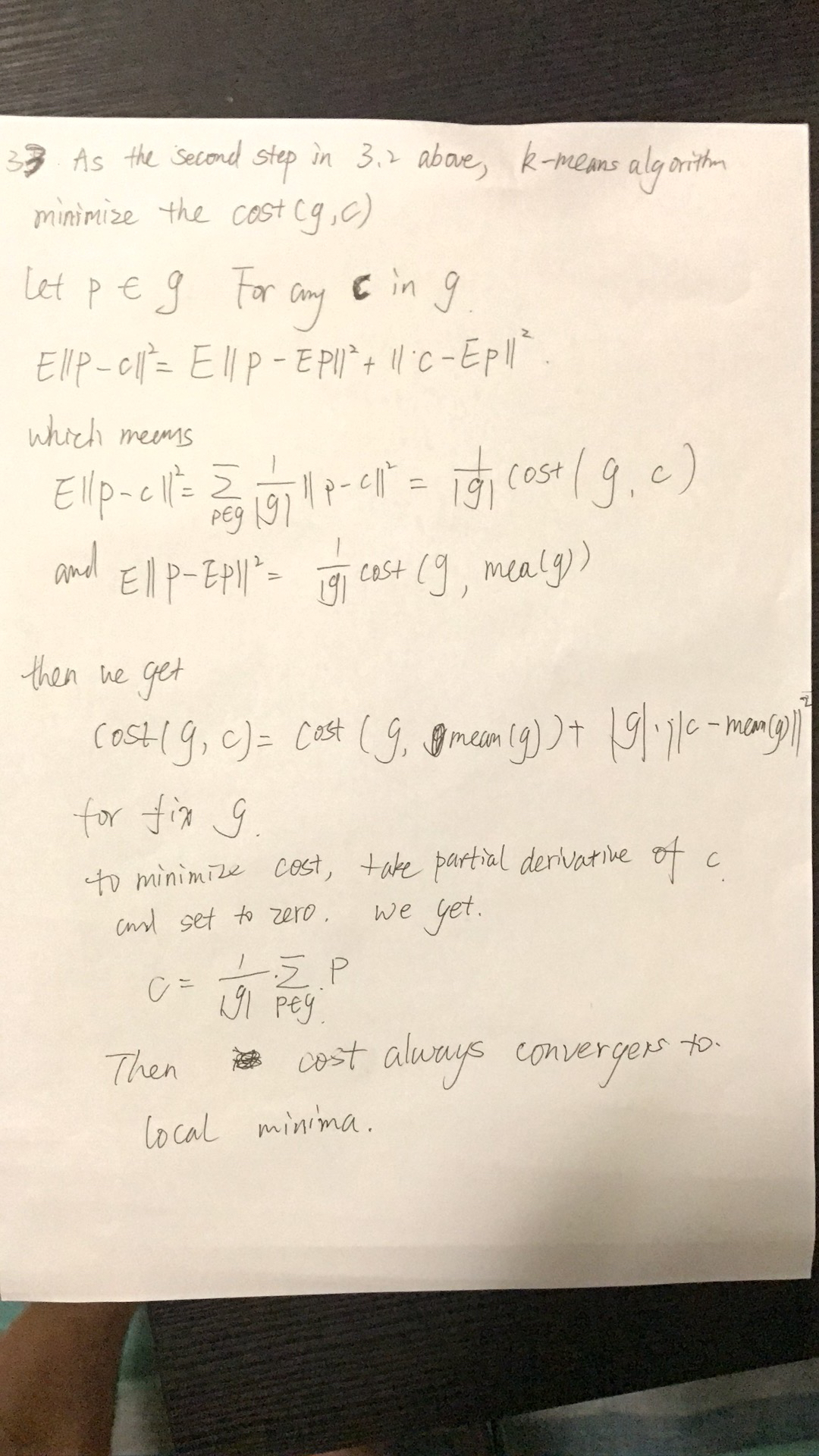
**C.append(ci);**

**if oldCentres == C then**

**canStop** ← **true**

10 return G;

next question is in next page

2)

3)