The Hong Kong Polytechnic University
Department of Computing
## COMP5511 – Introduction to Artificial Intelligence
Semester 1, 2017-18

Assignment 2
Due Date: November 15, 2018

1. (50 *marks*) The ABC Telecom has collected more demographic data from its customers. A sample of 15 such data records were selected and are shown below. Each record is characterized by six attributes: (i) Age; (ii) Sex, (iii) Monthly Income; (iv) Marital status; (v) Service Plan (in mins.) and (vi) Extra Usage (in mins.).

(0)	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Ref	Age	Sex	Monthly Income	Marital Status	Service Plan	Extra Usage
1	54	FEMALE	3000	YES	100	0
2	59	FEMALE	4000	NO	600	54
3	38	MALE	7800	NO	200	31
4	18	FEMALE	8500	NO	600	311
5	27	MALE	14000	YES	100	211
6	29	FEMALE	31000	YES	1600	25
7	17	MALE	7500	NO	600	254
8	22	FEMALE	7900	NO	200	31
9	34	MALE	24700	NO	100	7
10	46	FEMALE	31110	YES	600	0
11	39	FEMALE	21000	YES	800	64
12	35	FEMALE	30000	NO	1600	0
13	39	MALE	40500	YES	1600	50
14	18	MALE	7800	NO	1000	290
15	22	MALE	18000	YES	400	303

a) Find a clustering arrangement of records using the *k*-mode algorithm by setting k=3 and using the records 1, 8 and 15 as the initial cluster centers. For those numeric columns, divide the range into 3 intervals (use Equal-width and Equal-depth, i.e. you should have two sets of results). You must show the steps of the first iteration.

b) Perform clustering by using the *k*-means algorithm but this time you can use Python to help find the results.

c) Compare all the results you obtained with the different algorithms. Which one would you consider as the best? Why?

2. (50 *marks*) You are given the following graph representing ten individuals that are related to each other in some ways in a social network.
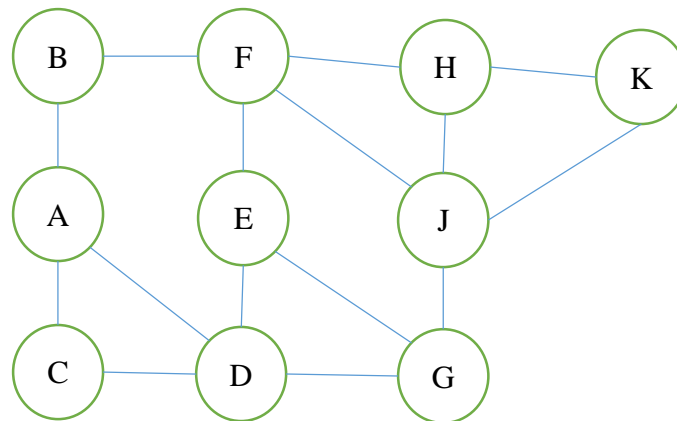
a) Use Breadth First Search or Depth First Search to find all the shortest paths between the ten different vertices in the graph. Show your work as samples and the results.
b) Compute the Edge-Betweenness values of the edges and put them in a table in the same way shown in the class notes. Show some samples of your work and the results.
c) Use the Girvan-Newman algorithm to discover in the graph two communities. Show your work and the results.