

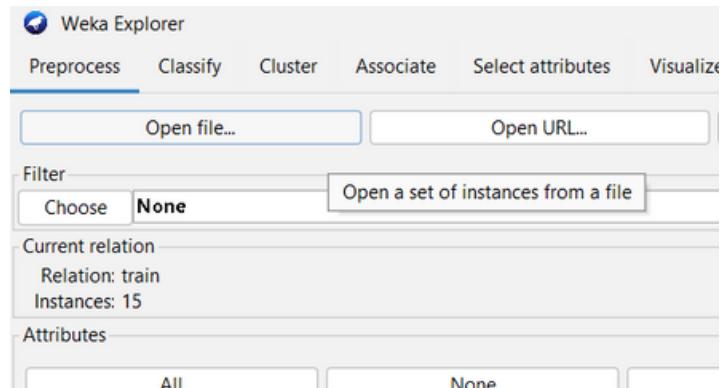
# ลองทำดู #1 : Clustering train data

1. เปิด Weka > เลือก Explorer > กดปุ่ม Open file ... > เลือกไฟล์ train.arff

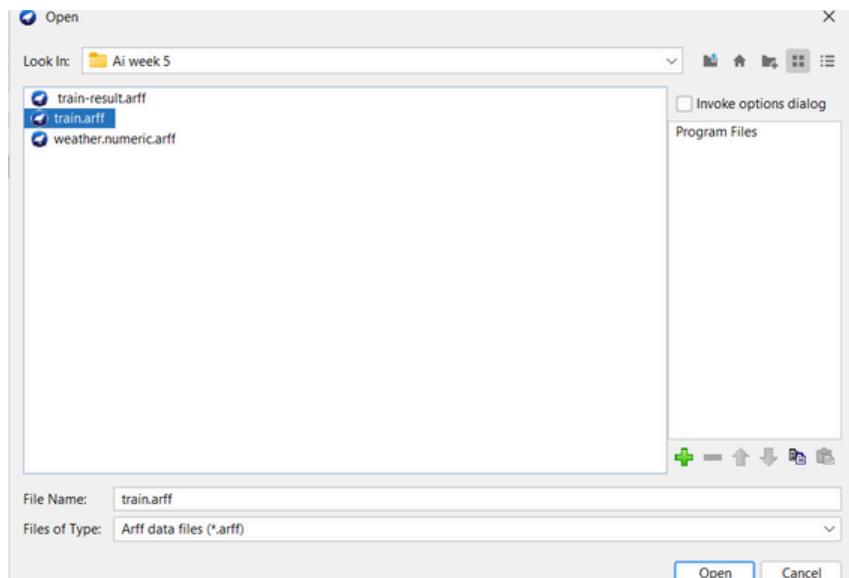


< 1. เลือก Explorer

2. จากนั้นกด Open file>  
<ด้านซ้ายบน>



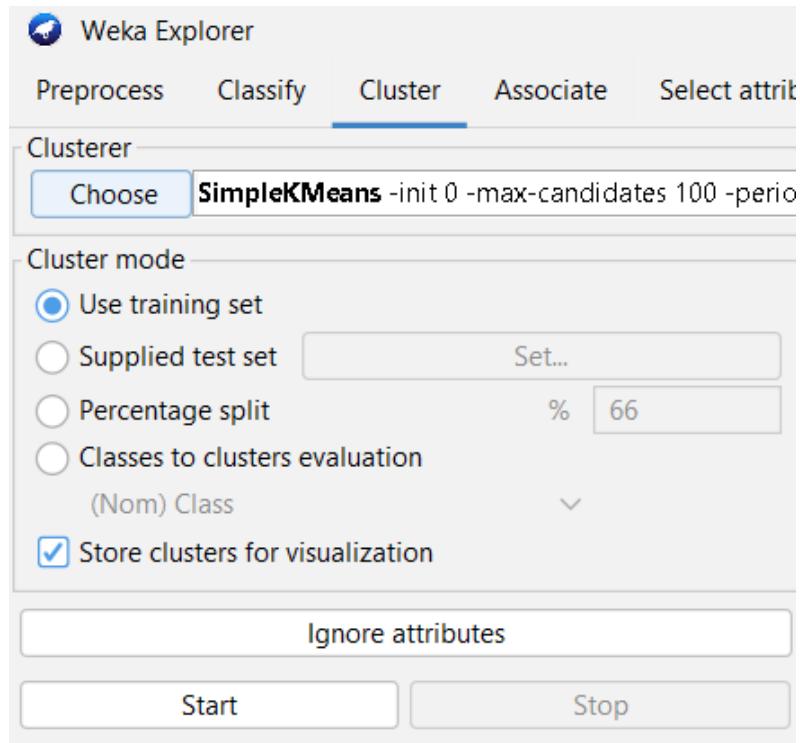
3. แล้วก็เลือกไฟล์ train.arff  
<หาไฟล์train.arff ที่saveไว้>



< ถ้าเลือกแล้วกด Open

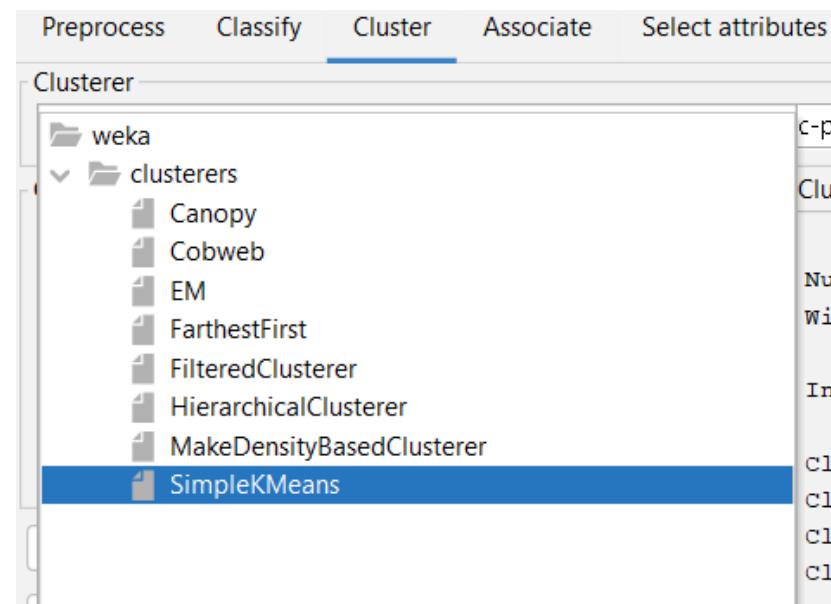
< 1. ต่อมาคลิกที่ tab Cluster  
<ด้านซ้ายบนของ WEKA>

2. จากนั้นกด Choose >



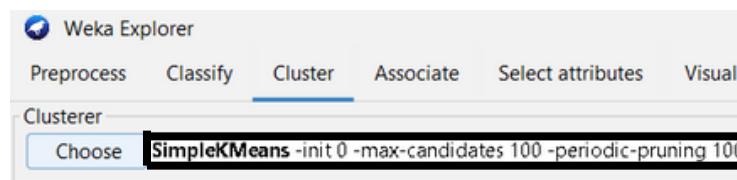
มันจะเด้งหน้านี้ขึ้นมาถ้ากด Choose

3. จากนั้นกด SimpleKMeans >

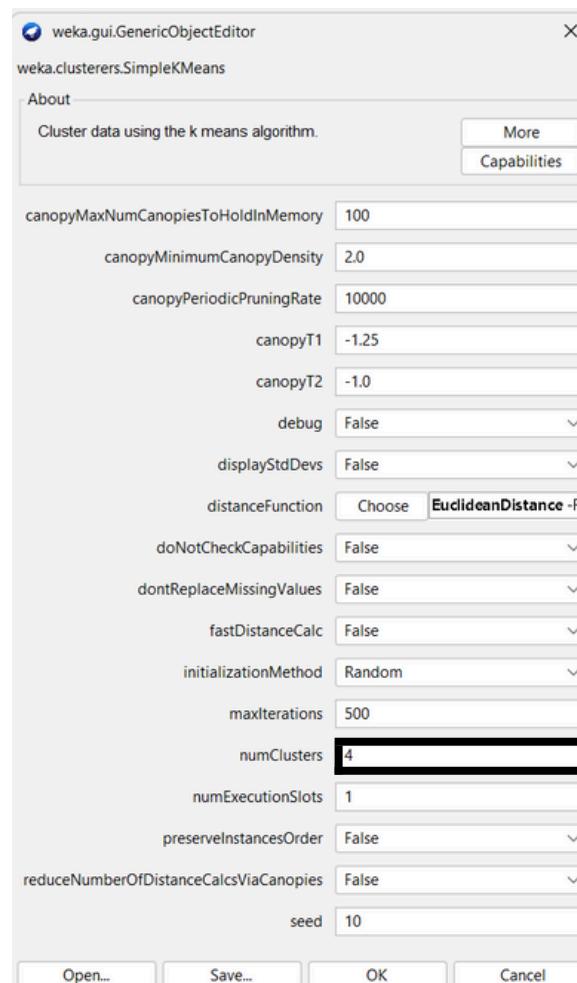


## โจทย์นอก เปลี่ยนค่าพารามิเตอร์ numClusters (บางครั้งจะเขียนย่อ ๆ ว่า N) เป็น 4

1. จักหนันคลิกซ้ายในช่องที่ ชื่อว่า SimpleKMeans >

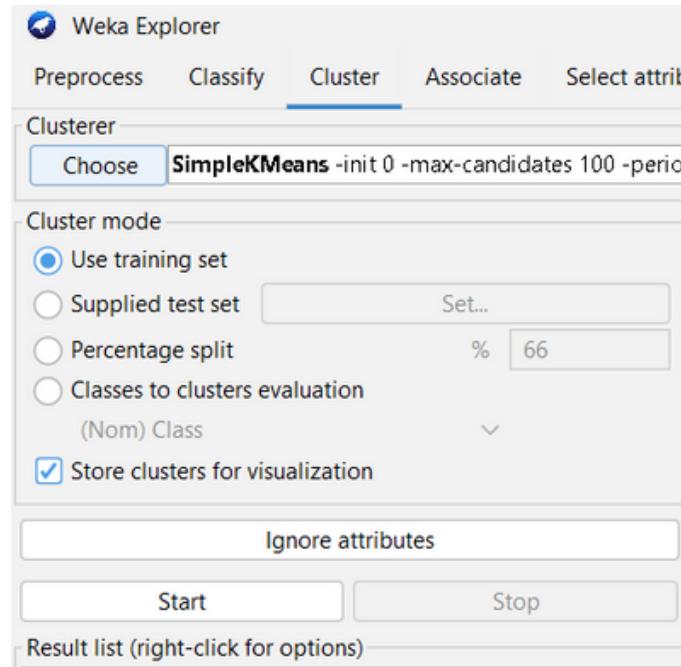


มันจะเด้งหน้านี้ขึ้นมา

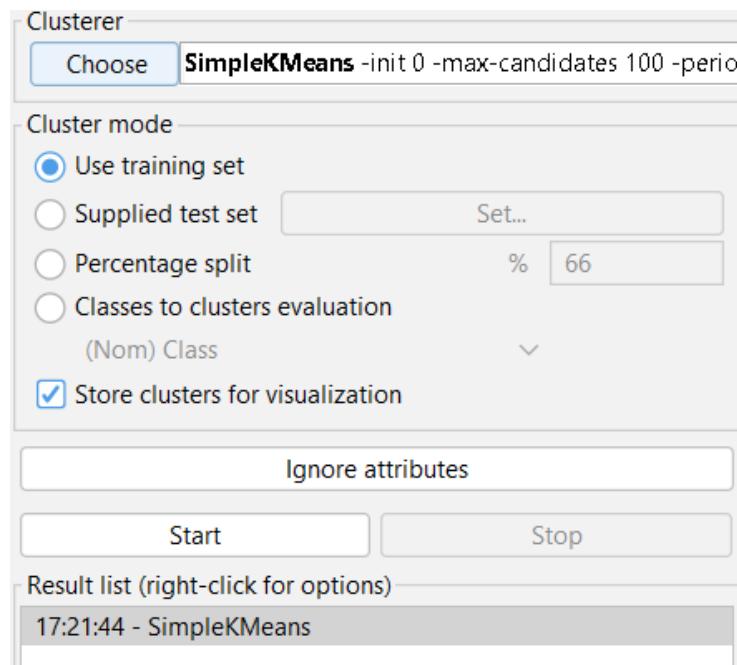


< 2. จักหนันกด numClusters เป็น 4

< 3. เสร็จแล้ว กด OK



1. คลิกปุ่ม Start >



,มันจะเด้งมาตรงนี้ถ้ากด Start >

# หน้าตาผลลัพท์

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Clusterer

Choose SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 4 -A "weka.core.EuclideanDistance -R first-last" -l 500 -num-slots 1 -S 10

Cluster mode

Use training set  
 Supplied test set Set...  
 Percentage split % 66  
 Classes to clusters evaluation (Nom) Class ▾  
 Store clusters for visualization

Ignore attributes

Start Stop

Result list (right-click for options)

17:21:44 - SimpleKMeans

Cluster output

```
Number of iterations: 3
Within cluster sum of squared errors: 18.698918560481786

Initial starting points (random):

Cluster 0: 4,Dang,M,25,82,A
Cluster 1: 9,Gaga,F,14,70,C
Cluster 2: 12,Rika,F,19,77,B
Cluster 3: 7,Sompong,M,18,78,B

Missing values globally replaced with mean/mode

Final cluster centroids:
```

Attribute	Full Data	0	1	2	3
	(15.0)	(1.0)	(5.0)	(4.0)	(5.0)
ID	8	4	10.2	11	4.2
Name	John	Dang	David	Sunee	John
SEX	F	M	F	F	M
Age	16.8667	25	12	24	14.4
Score	70.4	82	61.6	83	66.8
Class	B	A	C	A	B

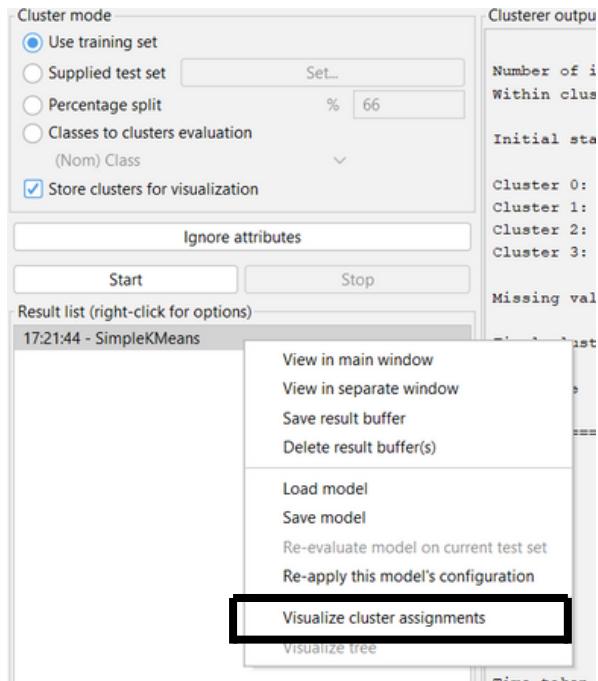
Time taken to build model (full training data) : 0 seconds

==== Model and evaluation on training set ===

Clustered Instances

```
0      1 ( 7%)
1      5 ( 33%)
2      4 ( 27%)
3      5 ( 33%)
```

1. คลิกขวาที่ผลลัพธ์ >



< 2. เลือก Visualize cluster assignment เพื่อแสดงกราฟ 3 มิติ

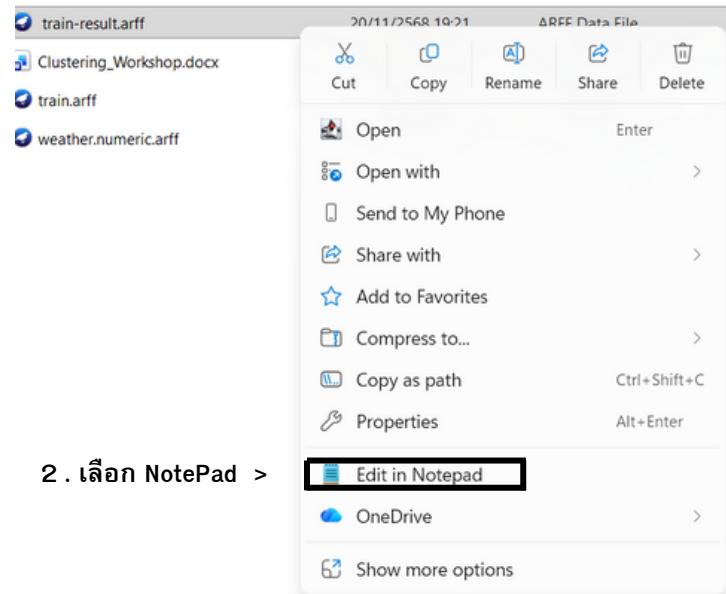
3. กดปุ่ม Save ในหน้าต่าง Weka Clusterer Visualize

คลิกแล้วหน้านี้จะเด้งขึ้นมา \*



## 1. Save ในชื่อ train-result.arff และทดลองเปิดดูในโปรแกรม WordPad หรือ NotePad

Name	Date modified	Type	Size
train-result.arff	20/11/2568 19:21	ARFF Data File	1 KB



2. เลือก NotePad >

จะเด้งขึ้นมาแบบนี้ >

```
File Edit View
@relation train_clustered

@attribute Instance_number numeric
@attribute ID numeric
@attribute Name {John,Somsak,David,Dang,Mana,Dum,Sompong,Sunee,Gaga,Lumyai,Momoko,Rika,Manee,Risa,Jessica}
@attribute SEX {M,F}
@attribute Age numeric
@attribute Score numeric
@attribute Class {B,C,A,D}
@attribute Cluster {cluster0,cluster1,cluster2,cluster3}

@data
0,1,John,M,17,71,B,cluster3
1,2,Somsak,M,18,75,B,cluster3
2,3,David,M,13,69,C,cluster1
3,4,Dang,M,25,82,A,cluster0
4,5,Mana,M,10,50,D,cluster3
5,6,Dum,M,9,60,D,cluster3
6,7,Sompong,M,18,78,B,cluster3
7,8,Sunee,F,26,85,A,cluster2
8,9,Gaga,F,14,70,C,cluster1
9,10,Lumyai,F,11,52,D,cluster1
10,11,Momoko,F,24,87,A,cluster2
11,12,Rika,F,19,77,B,cluster2
12,13,Manee,F,27,83,A,cluster2
13,14,Risa,F,14,62,C,cluster1
14,15,Jessica,F,8,55,D,cluster1
```

ถ้าเสร็จแล้วก็เอาไปตอบคำถามใน ปฏิบัติการรายงานปฏิบัติการ (Lab Report) ได้เลย