Customer Rating of Breakfast Cereals

The dataset Cereals.xls includes nutritional information, store display, and consumer ratings for 77 breakfast cereals, with 16 attributes

1. Name: Name of cereal											
2. mfr: Manufacturer of cereal where A = A	merican Ho	ome Food F	roducts; G	= General	Mills; K = K	elloggs; N =	Nabisco;	P = Post; Q	= Quaker	Oats; R = F	Ralston Purina
3. type: cold or hot						00 ,					
4. calories: calories per serving											
5. protein: grams of protein											
6. fat: grams of fat											
7. sodium: milligrams of sodium											
8. fiber: grams of dietary fiber											
9. carbo: grams of complex carbohydrates											
10. sugars: grams of sugars											
11. potass: milligrams of potassium											
12. vitamins: vitamins and minerals - 0, 25,	, or 100, in	dicating the	typical per	centage of	FDA recom	mended					
13. shelf: display shelf (1, 2, or 3, counting	from the fl	oor)									
14. weight: weight in ounces of one serving	j										
15. cups: number of cups in one serving											
16. rating: a rating of the cereals calculated	d by Consu	ımer Repor	ts								

<u>Data-preprocessing</u>: Remove all cereals with missing values, do normalization and appropriate attribute selection

Data Mining Steps:

- a) Apply K-Mean algorithm with best K
- b) Explain the meaning of each cluster by going their description.
- c) The elementary public schools would like to choose a set of cereals to include in their daily cafeterias. Every day a different cereal is offered, but all cereals should support a healthy diet. For this goal you are requested to find a cluster of "healthy cereals"
- Submit student_id_creal.pdf & student_id_creal.xml (or py)