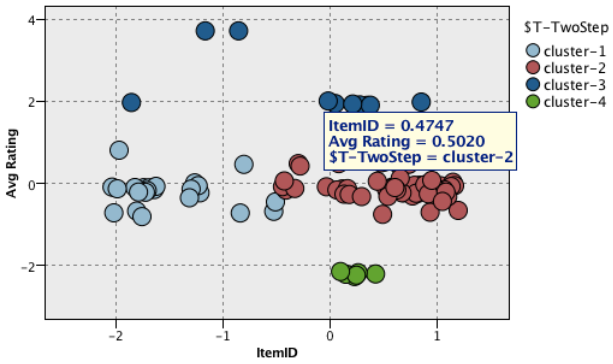
# SUMMARY

# PROBLEM 2

The Second Problem requires us to form combos between items having high average rating and low average rating as well as between items having low average rating, such that after decrease in price of the combo the loss in revenue is minimal.

The data mining techniques used to solve this problem are **Clustering followed by Association Rule Mining**.
**Two Step clustering** has been used to find clusters of items having high average rating and low average rating. To apply this clustering, first we need to compute the average rating of all items from the dataset given. A total of four clusters were formed. We get ten items in the cluster of high average rating and six in the cluster having low average rating.




Then we use **Apriori** to find the combos between the items in these two clusters. Apriori is an association rule mining technique that uses support based pruning to control the exponential growth of candidate item sets. The order of items in the combo, whether it is an antecedent or consequent, is not important.
After this, we eliminate redundant combos and the rules formed between items having high rating. We select the best ten combos because we need to find the minimum loss in revenue. The ten combos are:

1) Item128 + Item 141
2) Item 151 + Item 148
3) Item 128 + Item 134
4) Item 139 + Item 141
5) Item 139 + Item 128
6) Item 151 + Item 141
7) Item 139 + Item 134
8) Item 142 + Item 141
9) Item 146 + Item 141
10) Item 138 + Item 134

Calculating the new price of the combos and the decrease in revenue as result of it we get a **% 0.84** for the month of December.