

Training Data Set Creation

The training data set used for this analysis is shown below (first 5 lines):

sparkMLlibClustering.py				*clusterCenters_SC.txt		clusterCenter_v2.txt	
bound	method	NDFrame.head of	teamId	strength	currentLevel	price	
0	18	0.885470	1	421.0			
1	90	0.443162	1	388.0			
2	78	0.796923	1	176.0			
3	32	0.276723	1	141.0			
4	25	0.894529	1	328.0			
5	59	0.611500	1	644.0			
6	9	0.952176	1	513.0			
7	8	0.908269	1	382.0			
8	53	0.237000	1	677.0			
9	94	0.765549	1	450.0			
10	35	0.836061	1	710.0			

Dimensions of the training data set (rows x columns) : 44 x 4

of clusters created: 3