**Cluster Centers**

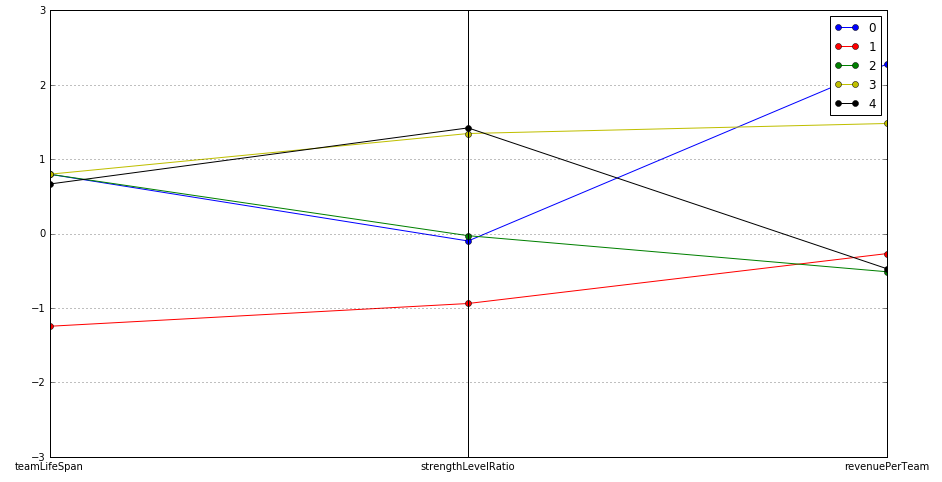
The code used in creating cluster centers is given below:



Cluster centers formed are given in the table below

|  |  |
| --- | --- |
| **Cluster #** | **Center** |
| 1 | 0.79811356, -0.09910393, 2.27760802 |
| 2 | -1.24554073, -0.93950776, -0.26919232 |
| 3 | 0.79811345, -0.02803835, -0.51249789 |
| 4 | 0.79811367, 1.34517521, 1.47997576 |
| 5 | 0.66626475, 1.42078479, -0.47074411 |

The difference between these clusters is visualized with the following graph:



1

2

3

4

5

These clusters can be differentiated from each other as follows:

**Cluster 1** is different from the others in that teams with an just above average teamLifeSpan and an average strengtLevelRatio have the highest revenuePerTeam.

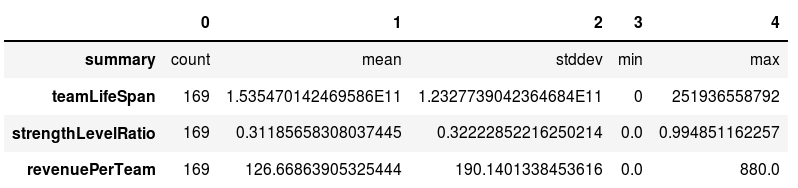
**Cluster 2** is different from the others in that the teams with a short teamLifeSpan (just started teams) and a low strengthLevelRatio have a just below average revenuePerTeam.

**Cluster 3** is different from the others in that the teams that are older than average and have an average strenghtLevelRatio have the lowest revenuePerTeam.

**Cluster 4** is what I hoped to find with this clustering. This cluster is different from others in that older teams (high teamLifeSpan) with a high strengthLevelRatio are willing to do in-game purchases (high revenuePerTeam).

**Cluster 5** is different from the others in that is contains old teams (high teamLifeSpan) with a high strengthLevelRatio and are not so willing to do in-game purchases (low revenuePerTeam).

Below you can see the summary of the train data set (not normalized):



print(km\_model.centers)

[array([ 0.79811356, -0.09910393, 2.27760802]),

array([-1.24554073, -0.93950776, -0.26919232]),

array([ 0.79811345, -0.02803835, -0.51249789]),

array([ 0.79811367, 1.34517521, 1.47997576]),

array([ 0.66626475, 1.42078479, -0.47074411])]