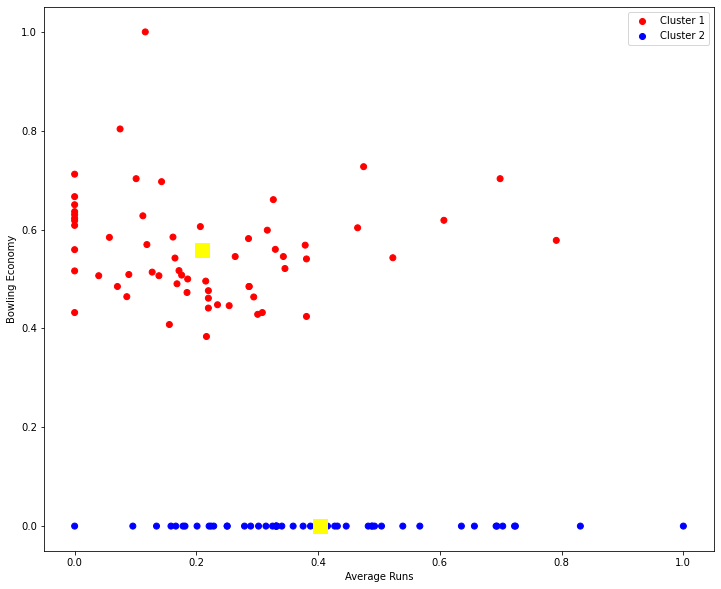
# **Unsupervised Learning - Assignment 1 [2019AIML622]**

**Question 2**

****

The above diagram shows the clustering information from cricket dataset.

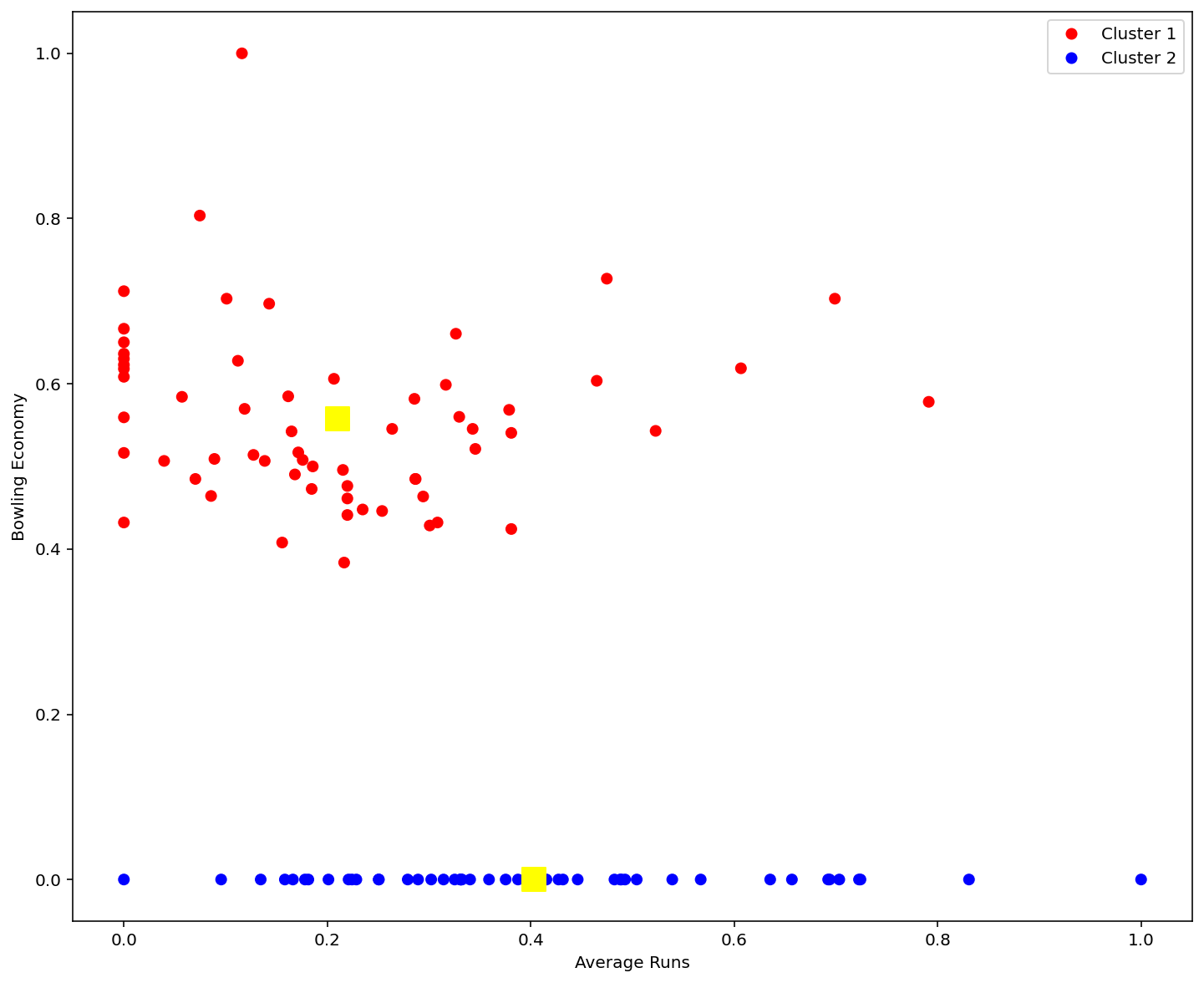
**Interpretation With K = 2 ( Two Clusters )**

* Cluster 1 (Red Color) → Bowling Economy is greater than Zero
* Cluster 2 (Blue Color) → Bowling Economy is equal to Zero
* From the visualization, though the bowling economy is zero we have average runs from 0 to 1. If team needs batsmen then we can choose from Cluster 2 with the average runs greater than 0.5
* If the team needs All Rounder then they can choose from Cluster 1 where average runs > 0.5 also if team needs bowler then they can choose from cluster 1

**Question 3**

**K = 2 ( Two Clusters )**

* Cluster 2 → Bowling Economy is Zero
* Cluster 1 → Bowling Economy is greater than Zero
* From the visualization, though the bowling economy is zero we have average runs from 0 to 1. If team needs batsmen then we can choose from Cluster 2 with the average runs greater than 0.5
* Similarly if the team needs All Rounder then they can choose from Cluster 1 with average runs > 0.5

****

**K = 3 ( Three Custers )**

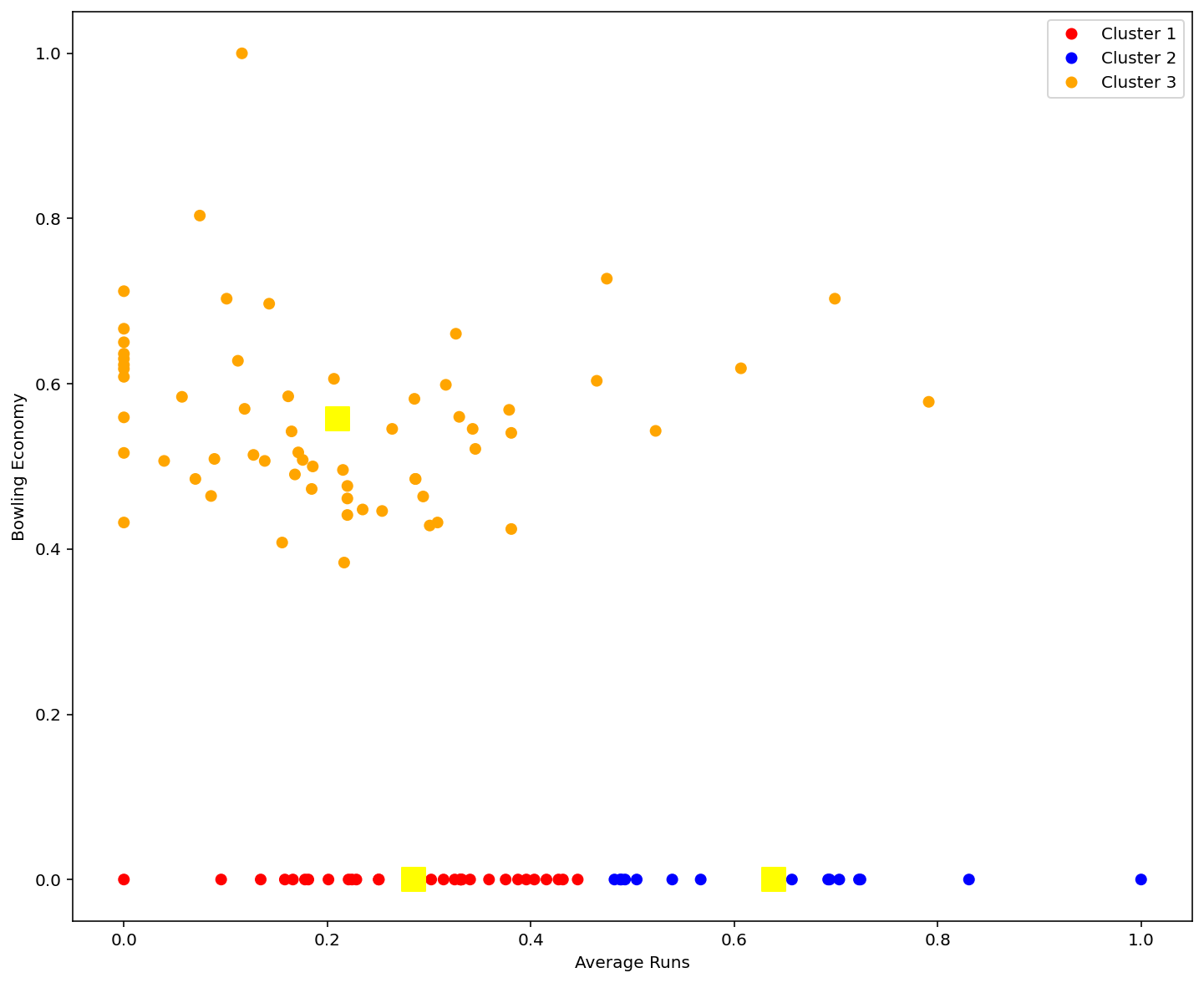
● Clusters 3 → Average Runs less than 0.5 and Bowling Economy Zero

● Cluster 2 → Average runs greater than 0.5 and Bowling Economy Zero

● Cluster 1 is formed with Bowling Economy greater than Zero

● If team needs batsmen then they can choose from Cluster 2

● If team needs All Rounder they can choose from Cluster 1 with average runs > 0.5

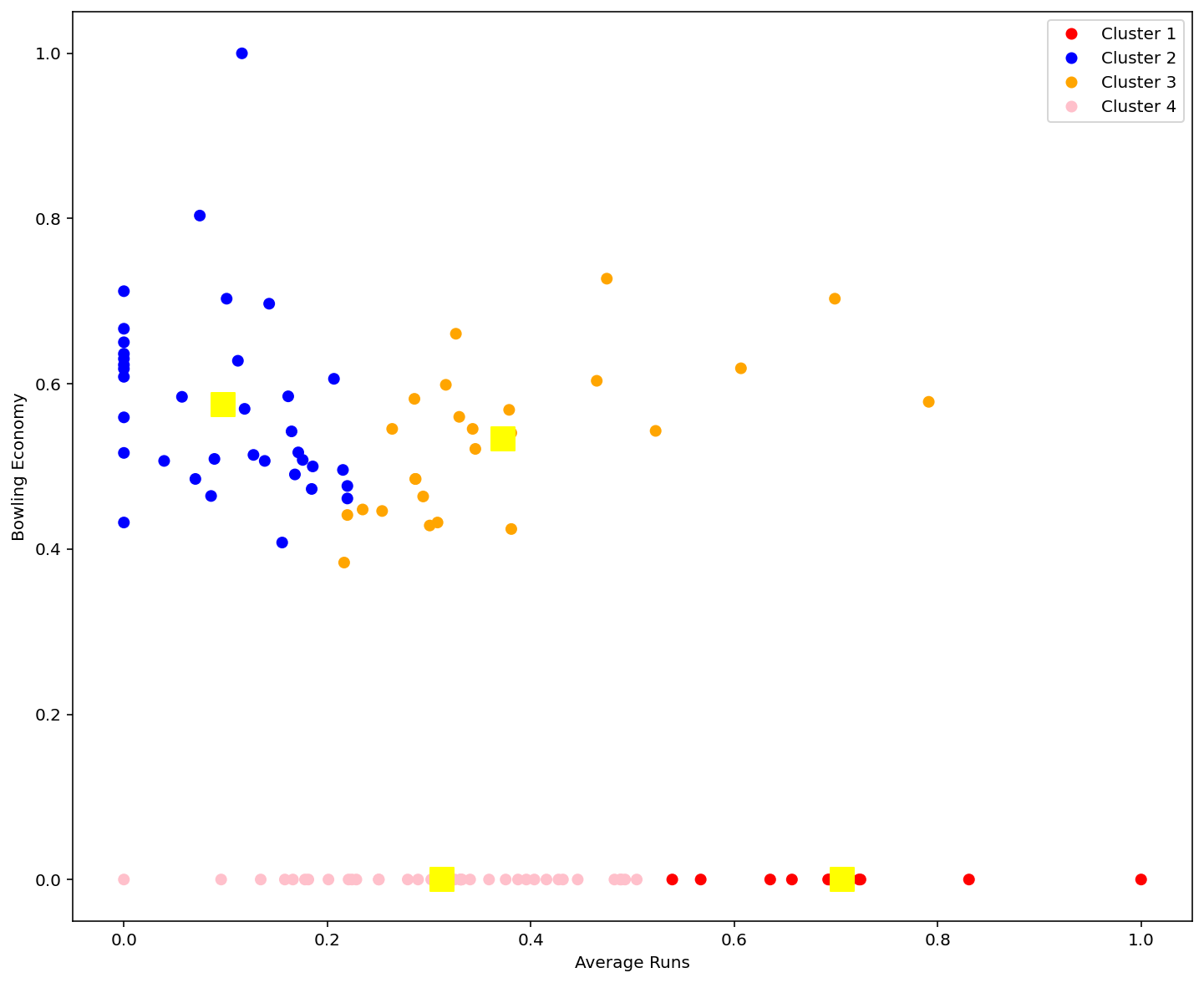


**K = 4 ( Four Clusters )**

● From the visualization if the team needs All rounder with high batting average and high bowling economy then they can choose from Cluster 4

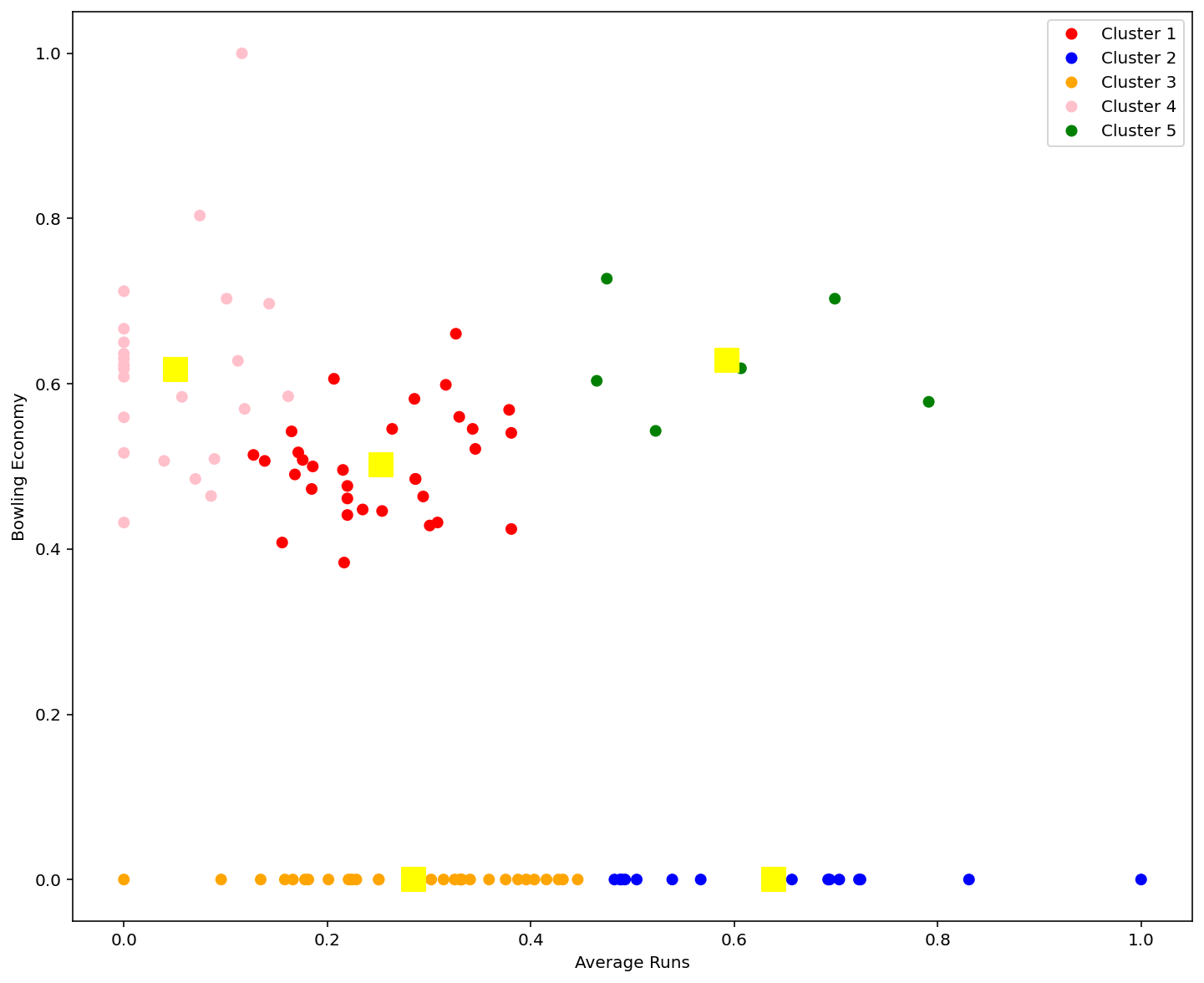
● If team needs high bowling economy then they can choose from Cluster 2

● If team needs batsmen then they can choose from cluster 3



**With K = 5, we got better clustering due to the following reasons**

* Cluster 1 → Average Runs greater than 0.5 and Bowling Economy Zero
* Cluster 4 → Average runs greater than 0.4 and Bowling Economy &gt; 0.4
* Cluster 5 is formed with Bowling Economy greater than 0.4
* If team needs batsmen then they can choose from Cluster 1
* If team needs All Rounder they can choose from Cluster 4
* If team needs bowler, they can choose from Cluster 3



**Question 4**

**Best Cluster**

* With K = 5 ( Five Clusters ) we got good clustering
* From the players list, we can group by Batsmen, Bowler, All Rounder
* Cluster 4 → Good batting and bowling
* Cluster 5 → Standard batting average
* Cluster 2 → Low batting average
* Cluster 3 → High bowling average
* Cluster 1 → High batting Average

**How can these clusters be useful?**

* If team needs All Rounder then they can choose from Cluster 4
* If team needs good bowler then they can choose from Cluster 3
* If team needs batsmen then they can choose from Cluster 1
* Cluster 2 are low batting average players, so they need good training in batting and to be focussed.