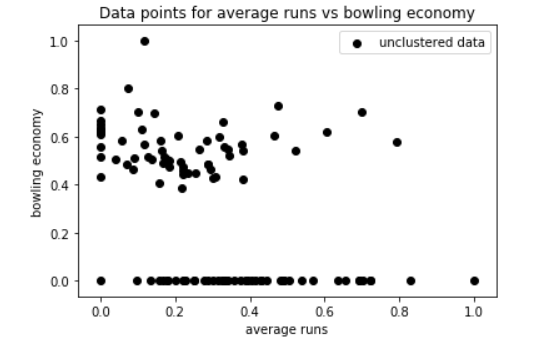
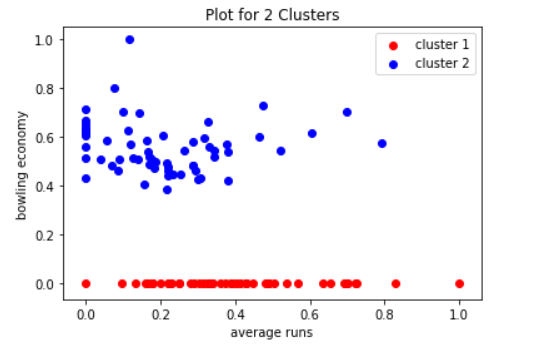
## Ans2.

The Data for cricketers’ career history is plotted for average runs vs bowling economy.



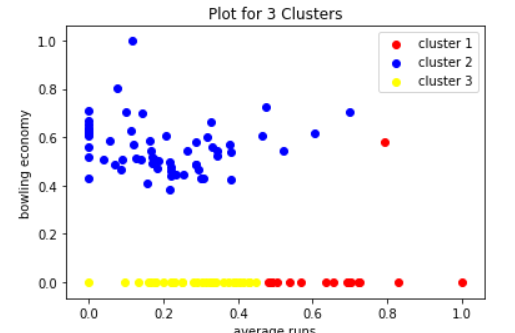
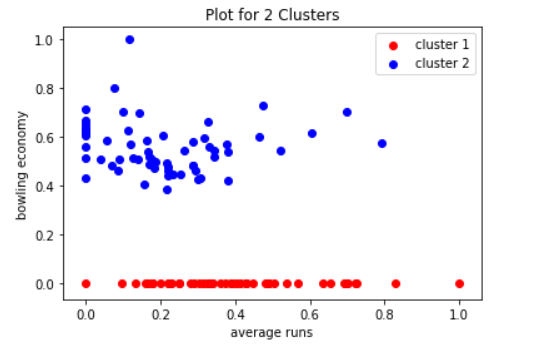
When the above data is clustered in 2 clusters using K-means algorithm then following clusters are obtained:

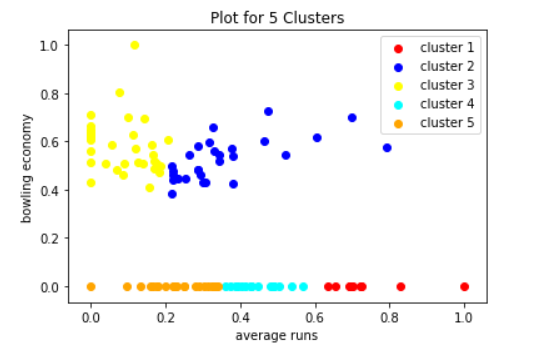
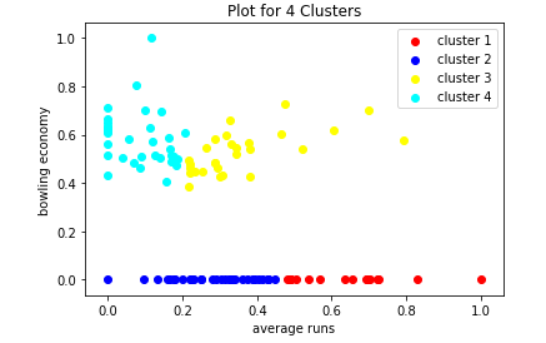


As per the clusters obtained above, we can clearly see the cricketers can be divided in two groups as bowlers and batsmen. Players with high Bowling economy have low average run, it states that they are full time bowlers. They come down in order and mainly do bowling. The Players which are highlighted in red have no bowling economy. That means they don’t bowl and are full time batsmen

## Ans3.

When the data points are clustered for k= 2, 3, 4 & 5, we get following charts:





Based on the above plot diagrams, the plot with 5 clusters seems to be best way to categorize players.

The players highlighted in yellow are the players which have very low batting average but are full time bowlers. Their high bowling economy states the fact that they are not good bowlers either. The economy should be low for a bowler.

The players highlighted in blue appear to be the players who are all-rounders. They have high bowling economy but they also have good number of average runs. So, they are fit for the all-rounder positions.

The players highlighted in orange are the batsmen who are not in form and they don’t bowl either. So, they might be sidelined for the selection.

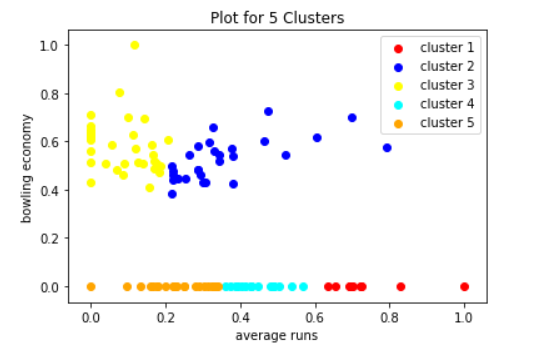
The players highlighted in red are the batsmen who have high average runs and almost nil bowling economy. So, this means they don’t bowl, are specifically batsmen and they seem to be good one too.

The players highlighted in cyan are also batsmen with nil bowling economy. These players have moderate average runs. They can be a good fit for middle order and can be taken as extras in tournaments.

The other plots do not divide players in this kind of detailed categories. Like in clusters 2 and 3, we cannot divide players in bowlers and all-rounders. In cluster 4, the players are divided but we cannot identify average batsmen who can be taken for middle order or need more chance to prove themselves and improve their game.

## Ans4.

The best plot seems to be the plot with 5 clusters.



It divides players in 5 categories: Bowlers, all-rounders, middle-order batsmen, good or opening batsmen and batsmen who are not in form since the data is being collected.

These categories can be used to identify team for coming tournaments. As, we have different category of players and their performance measured in terms of bowling economy and average runs scored.

This plot can also be used to divide players and then decide the contract they can be offered. Like the top batsmen (red), all-rounders (blue) can be given highest contract. If one is selecting players for tournaments like IPL, then he can decide the price cap for a particular based on his budget and player category.