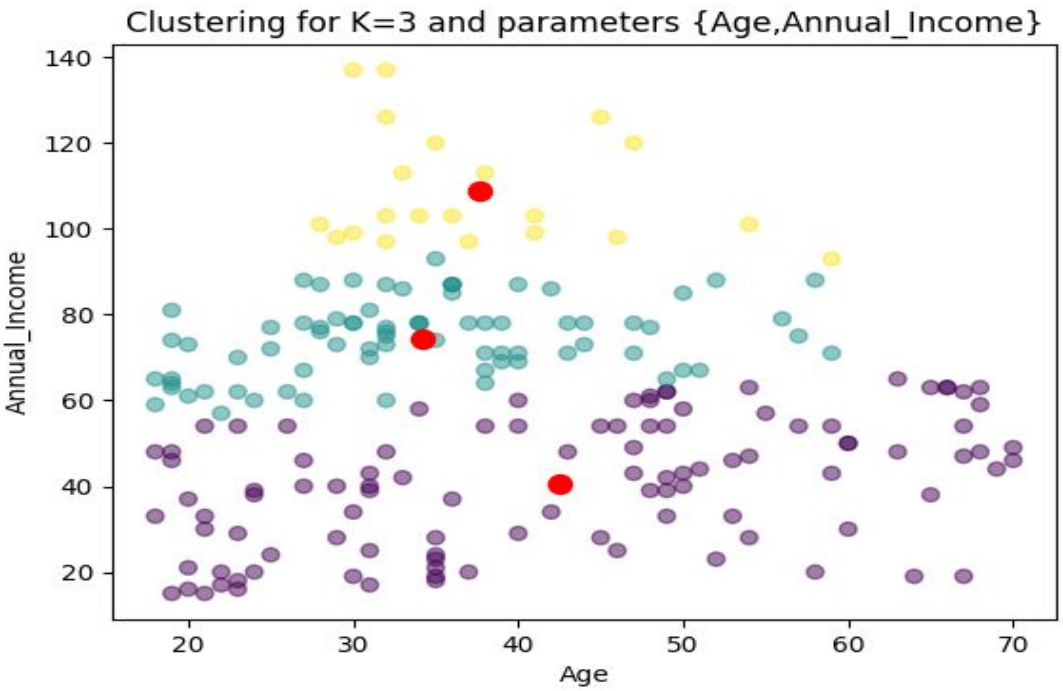
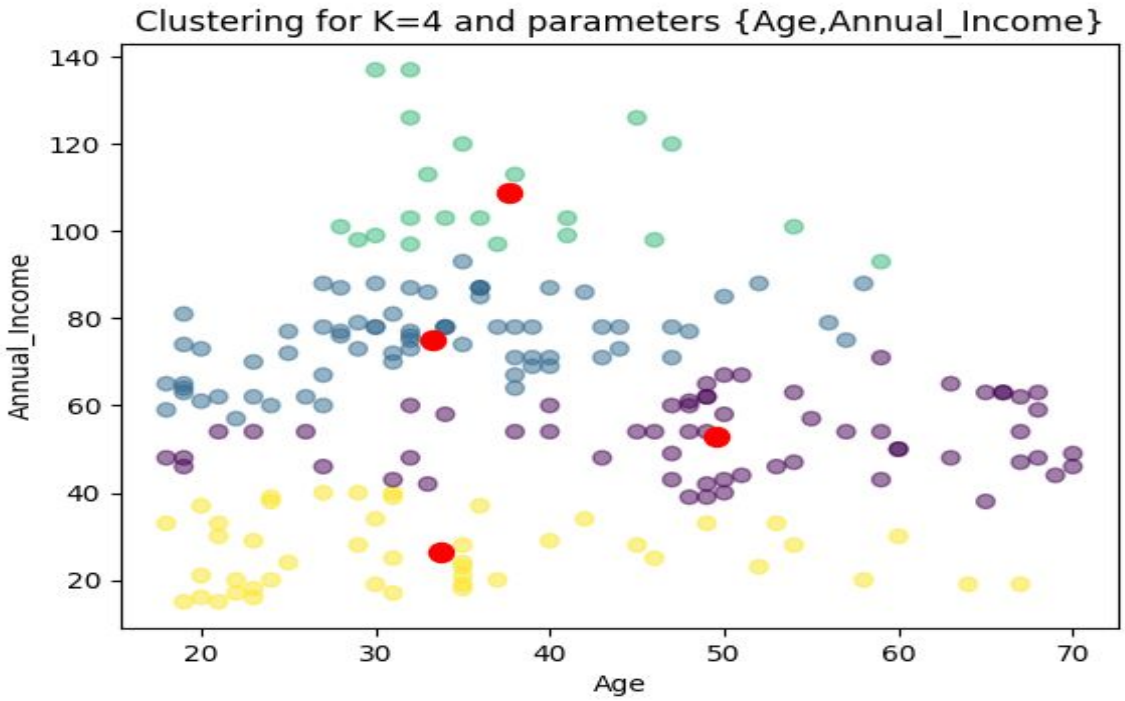
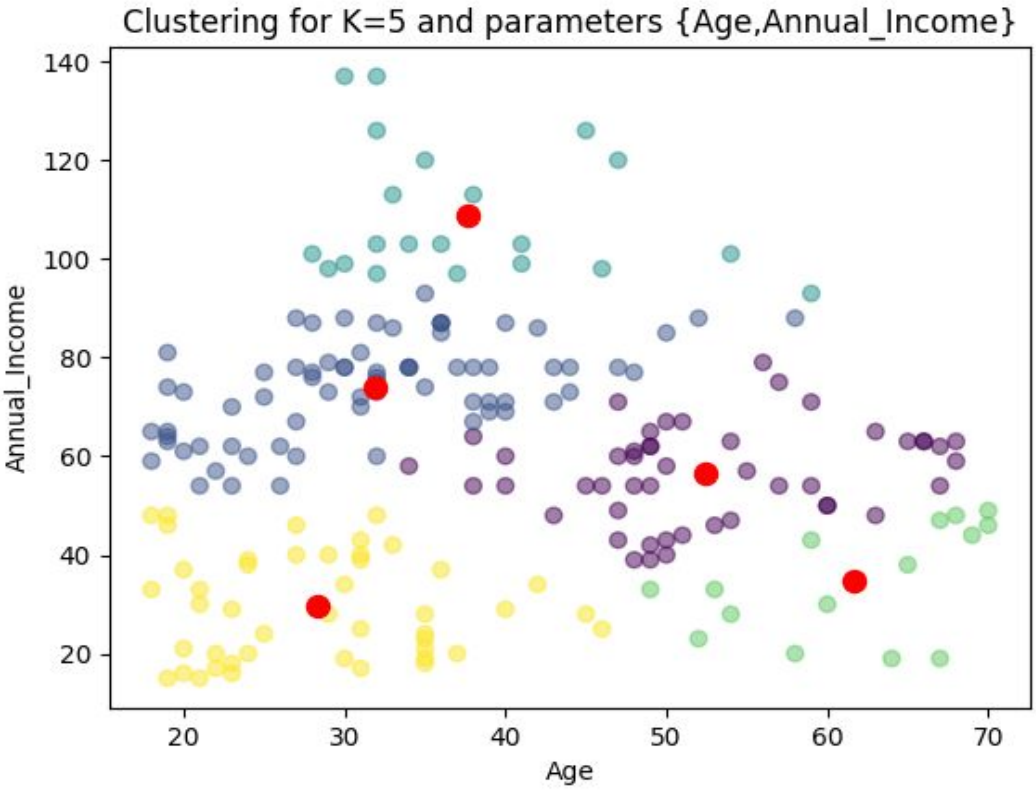
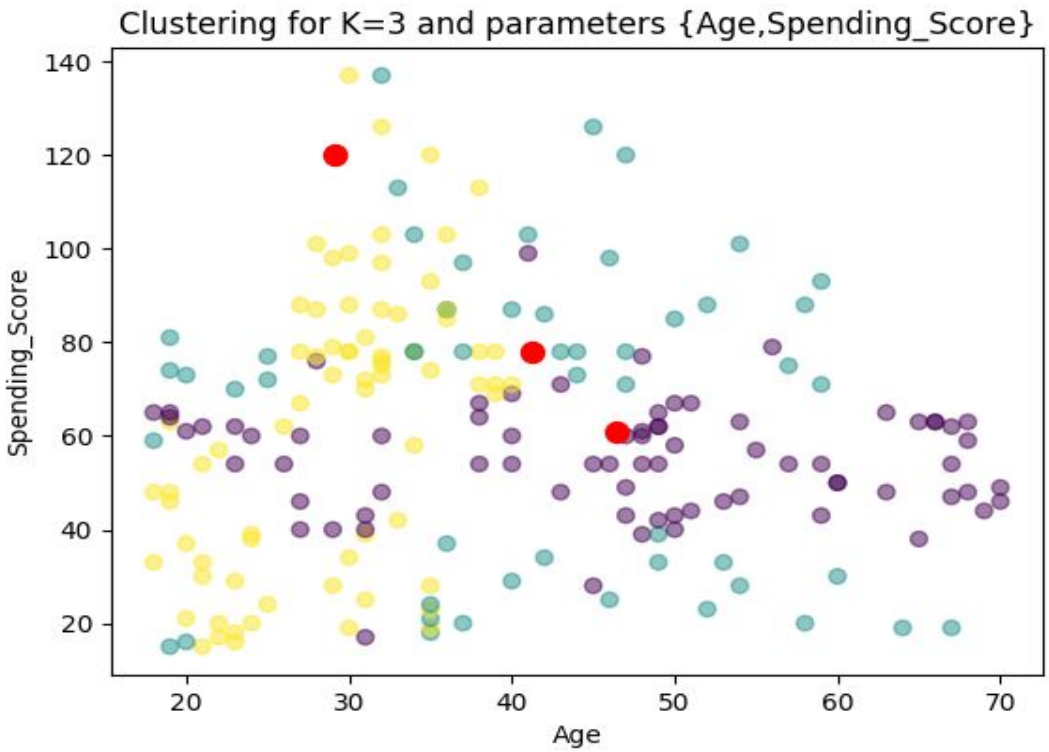
Clustering Coding Assignment Results:

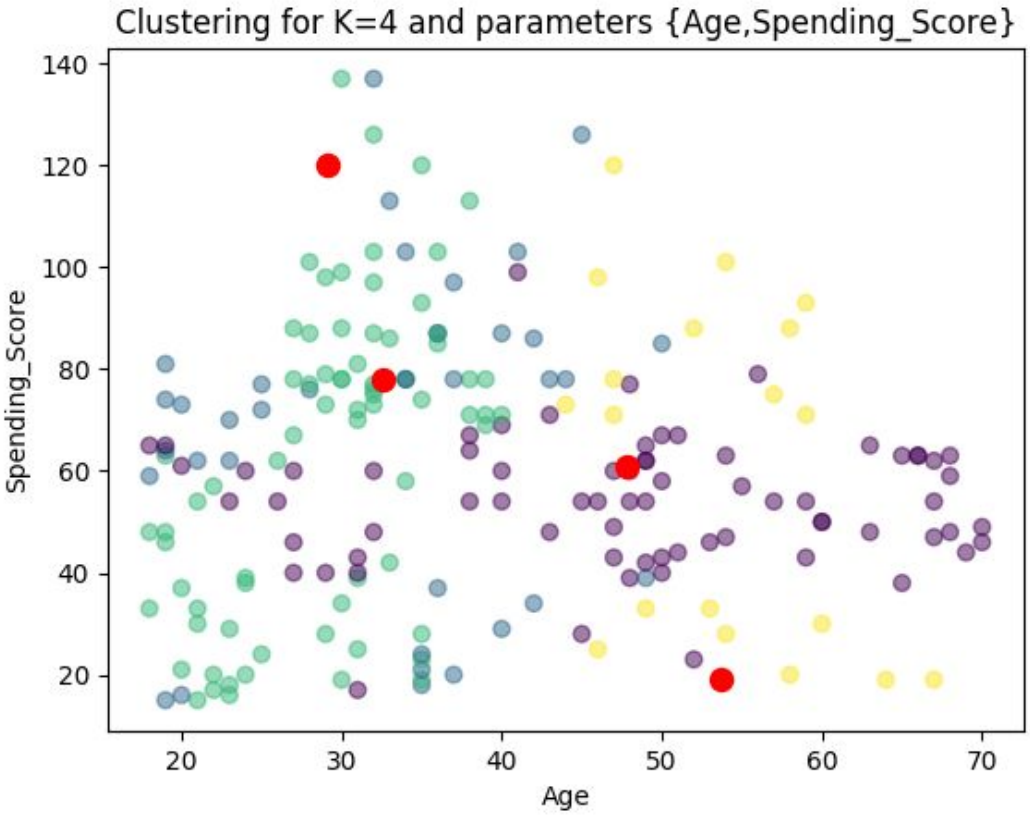
The resulting plots are:

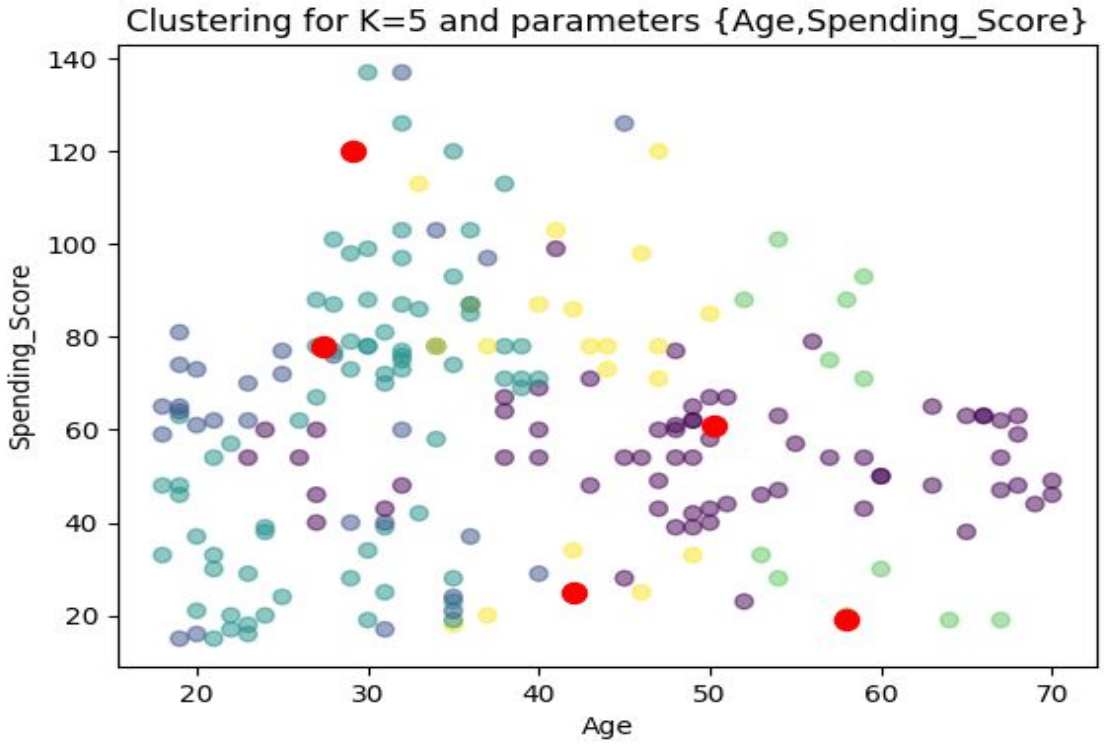


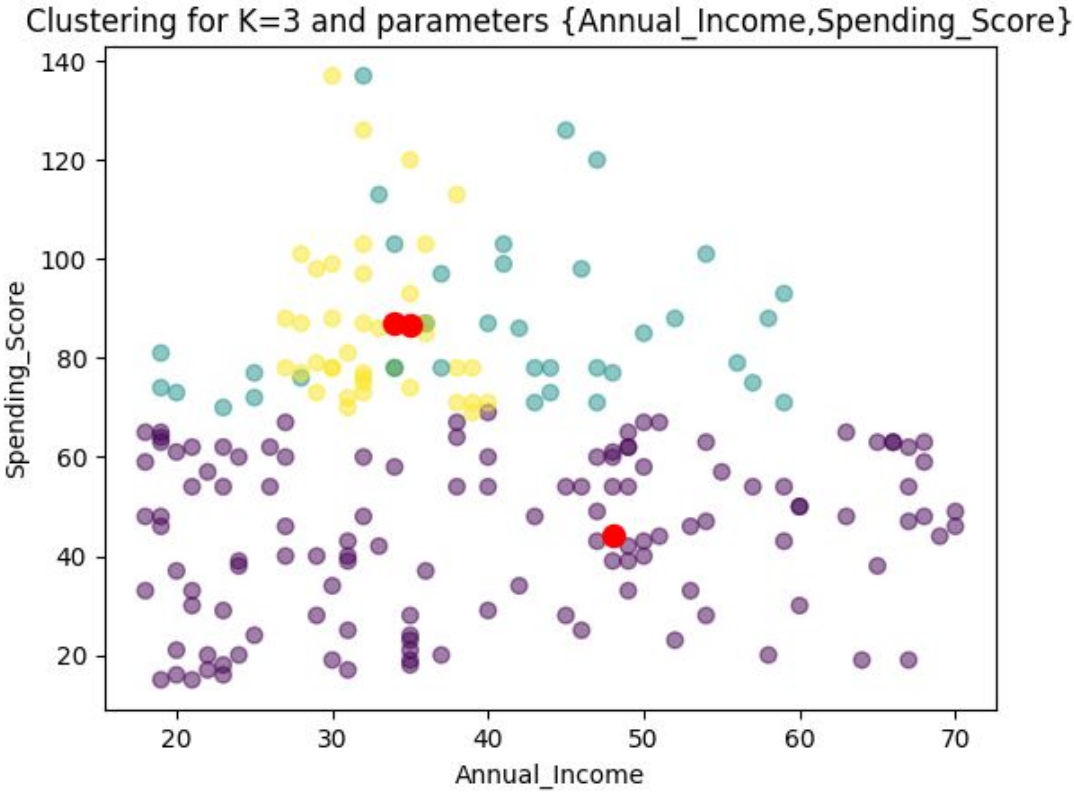


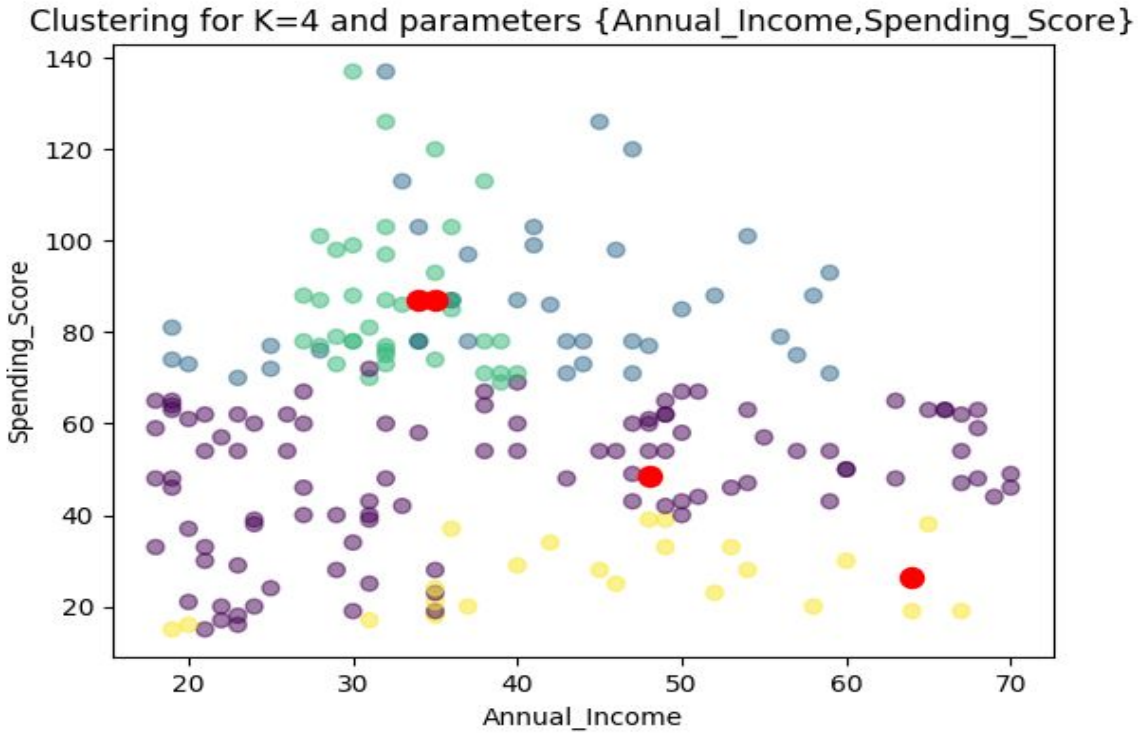

Clustering for K=5 and parameters {Age,Annual_Income}
Annual_Income
Age
140
120
100
80
60
40
20
20
30
40
50
60
70

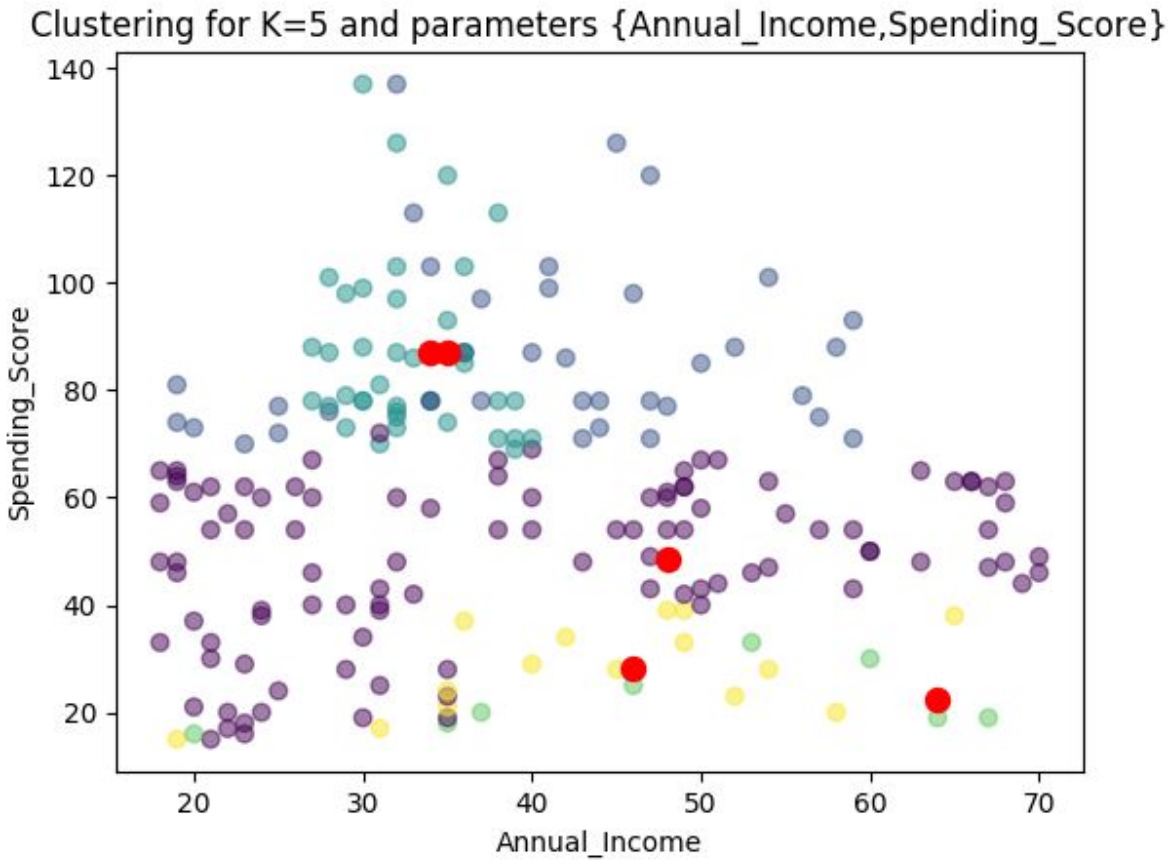

Clustering for K=3 and parameters {Age,Spending_Score}
Spending_Score
Age
140
120
100
80
60
40
20
20
30
40
50
60
70


Clustering for K=4 and parameters {Age,Spending_Score}
Spending_Score
Age
140
120
100
80
60
40
20
20
30
40
50
60
70


Clustering for K=5 and parameters {Age,Spending_Score}
Spending_Score
Age
140
120
100
80
60
40
20
20
30
40
50
60
70


Clustering for K=3 and parameters {Annual_Income,Spending_Score}
Spending_Score
Annual_Income
140
120
100
80
60
40
20
20
30
40
50
60
70


Clustering for K=4 and parameters {Annual_Income,Spending_Score}
Spending_Score
Annual_Income
140
120
100
80
60
40
20
20
30
40
50
60
70



The 3 sets of parameters {Age, Annual Income}, {Annual Income, Spending Score}, {Age, Spending Score} were used to cluster the given data using K-means clustering algorithm with K = 3,4,5 for each set of parameters.

Observations:

1. Best set of parameters to cluster the given dataset was {Age, Annual Income}. For these parameters, the best clustering was observed to be for K = 5.
2. Clustering using {Annual Income, Spending Score} resulted in poorly generated clusters with few of the centroids mixed up.
3. In case of parameters {Age, Spending Score} parameters, the best clustering output was observed in case of K=4.