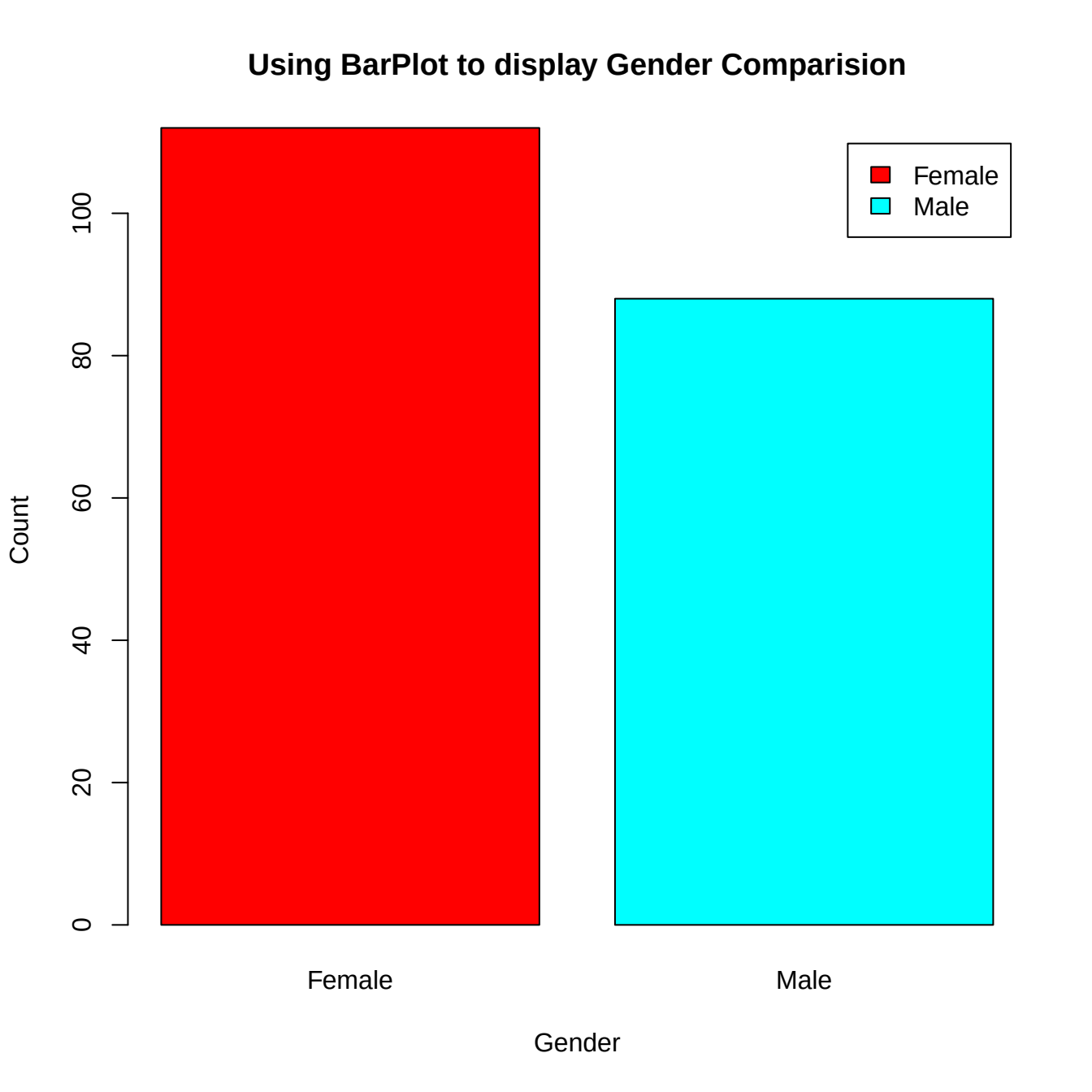
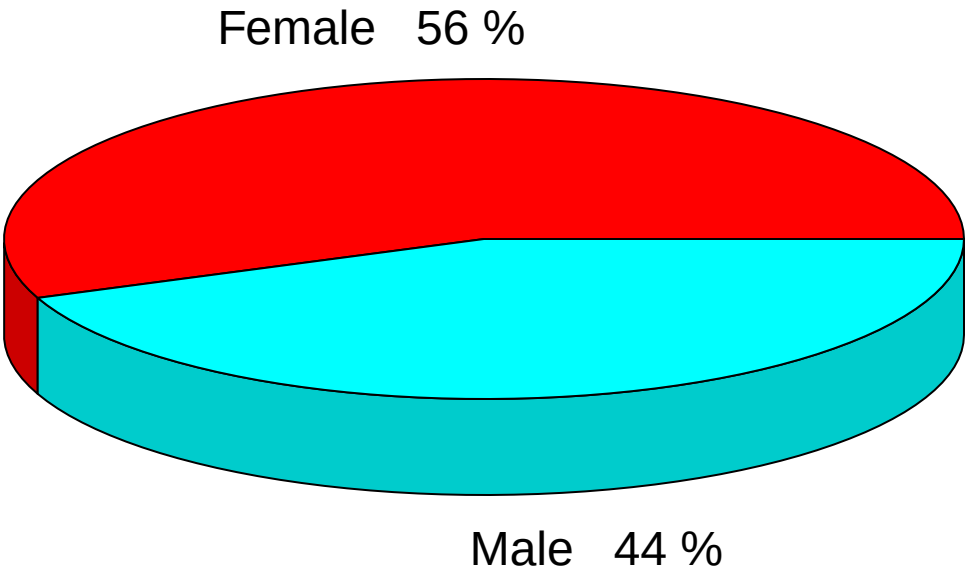
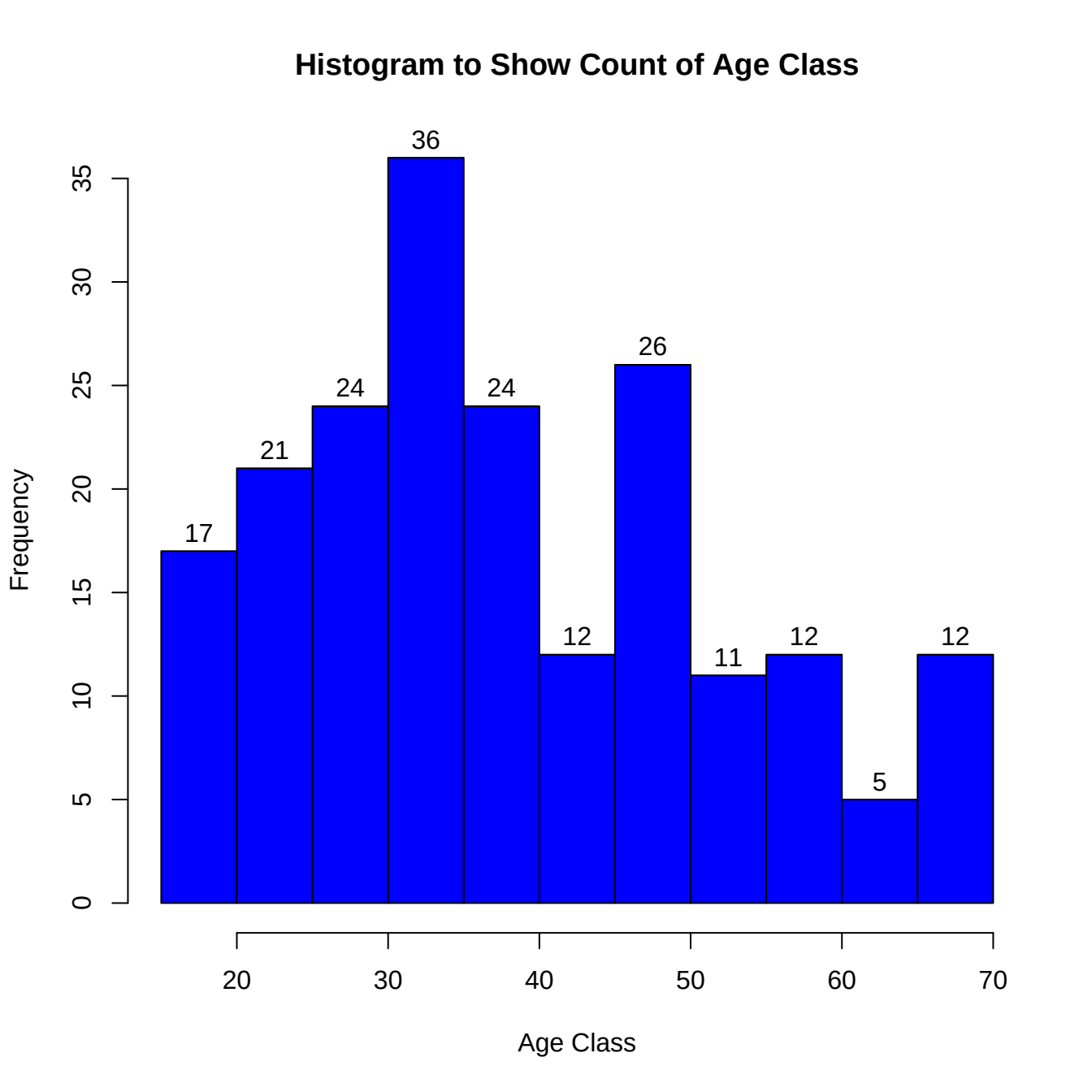
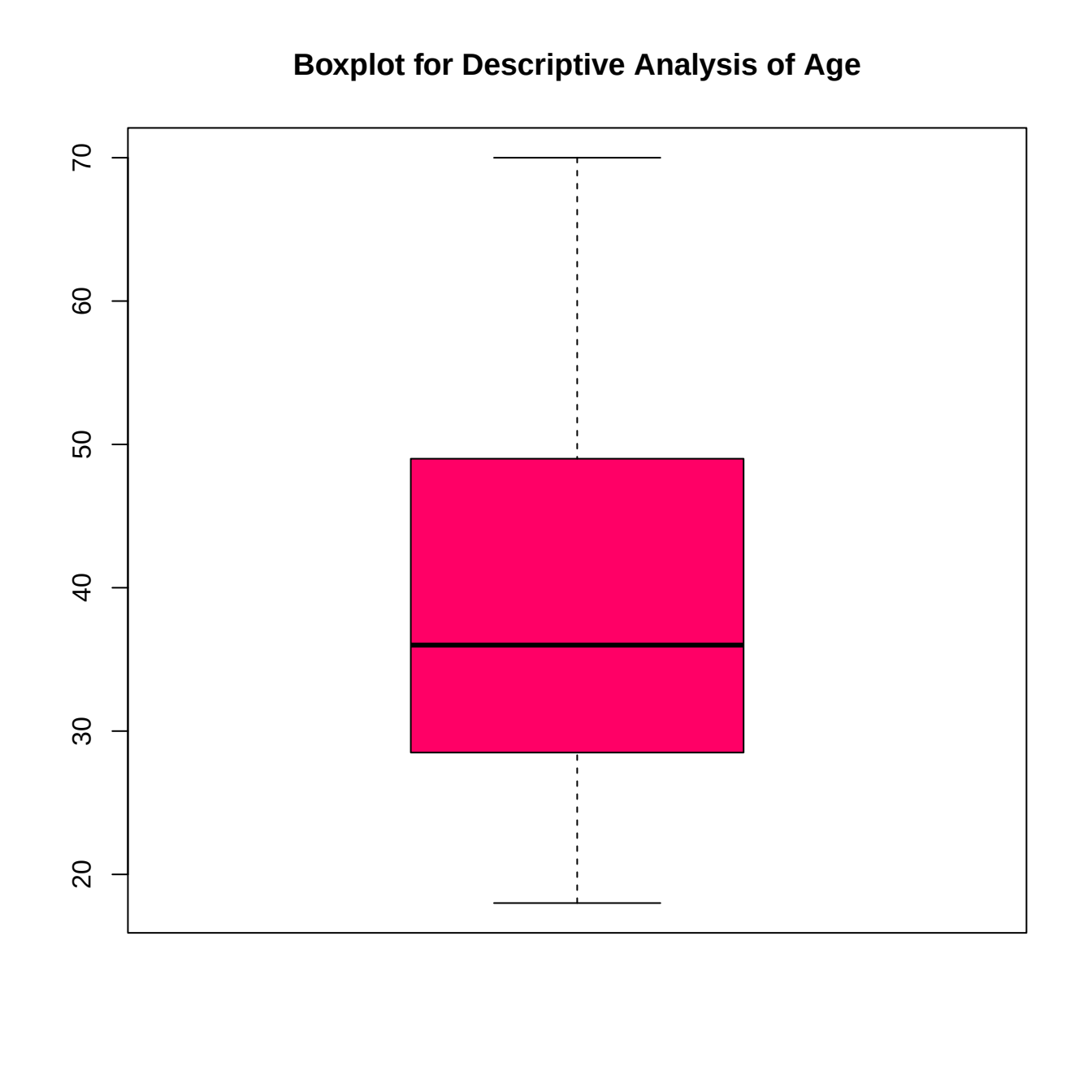
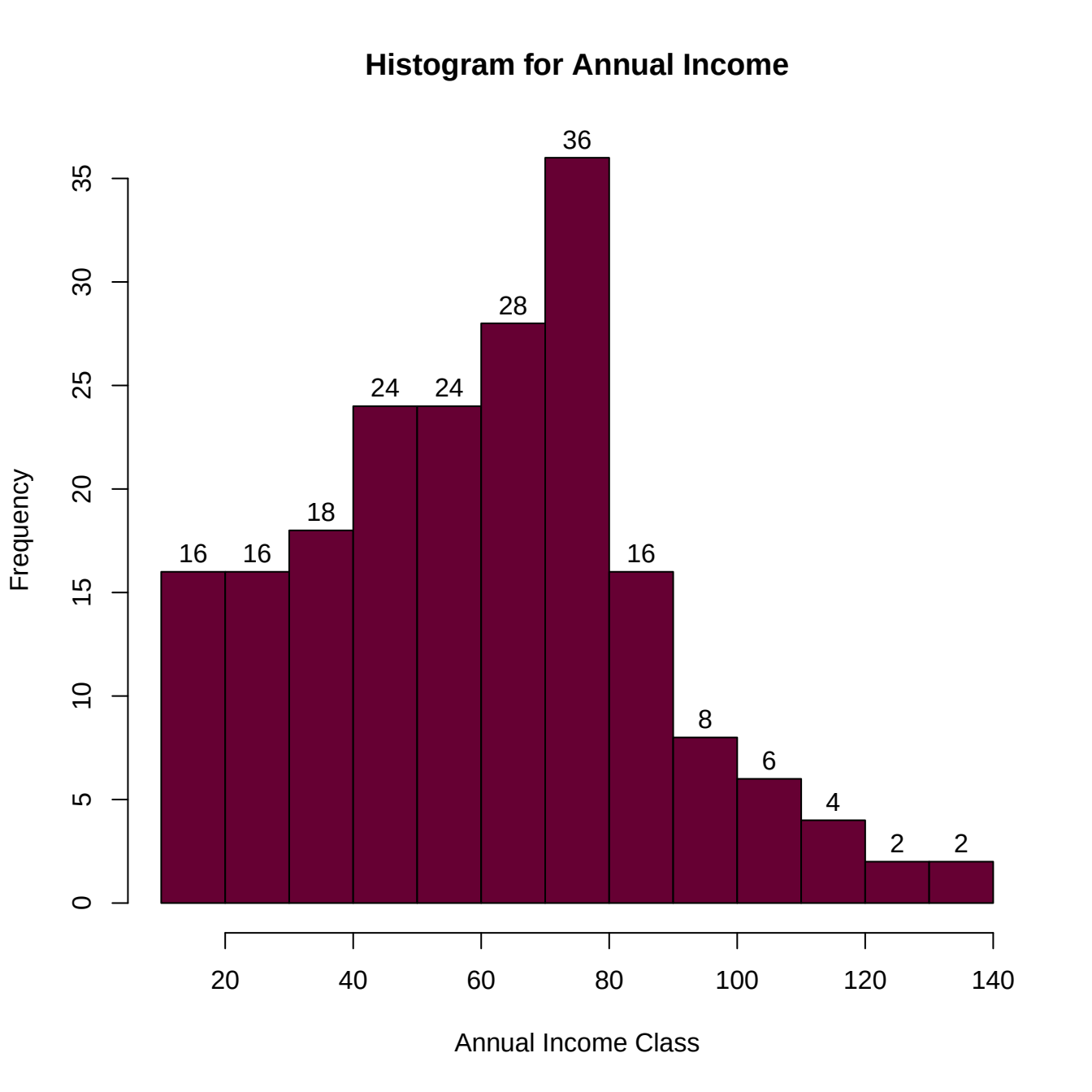

Using BarPlot to display Gender Comparision
Female
Male
Count
0
20
40
60
80
100
Female
Male
Gender

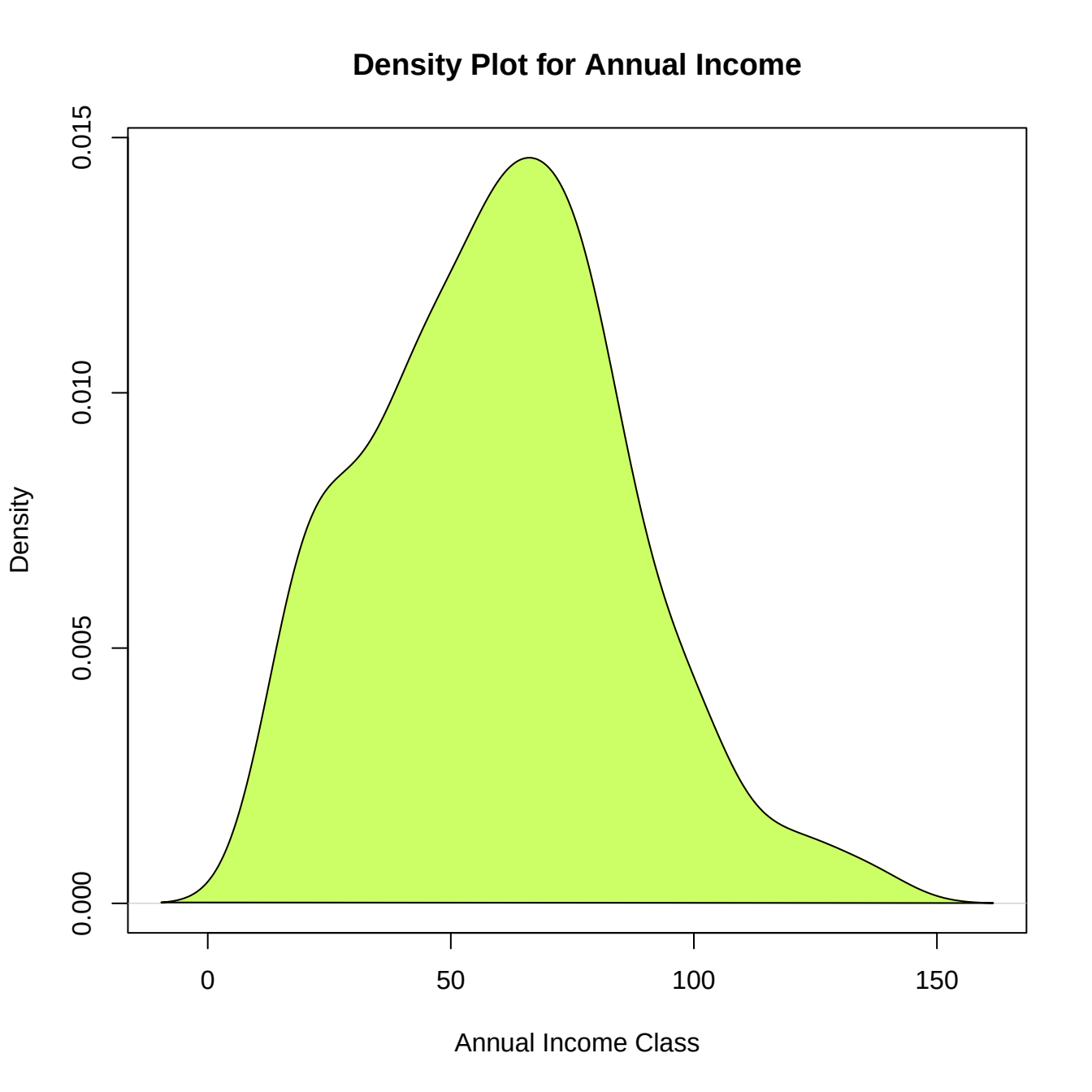
Pie Chart Depicting Ratio of Female and Male

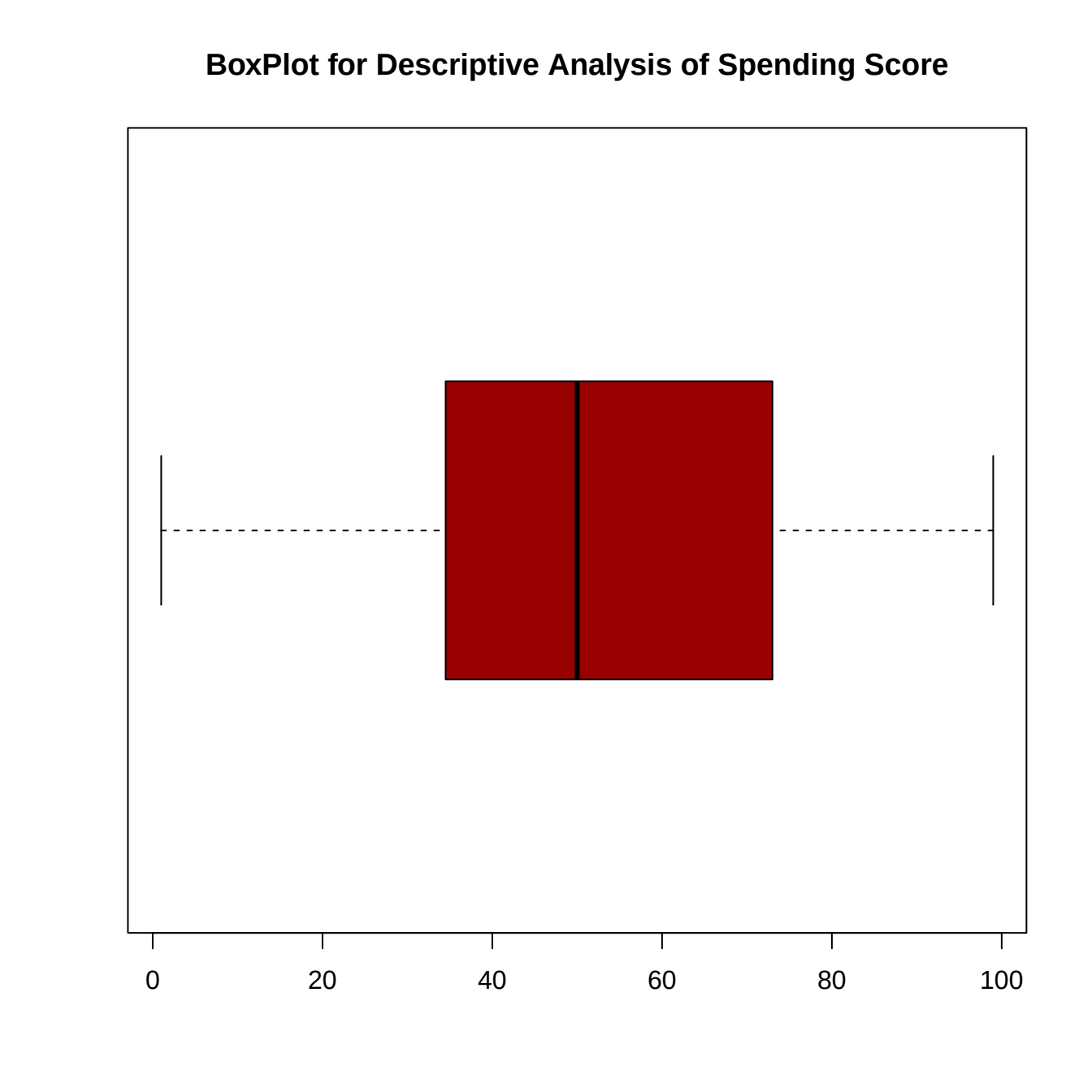


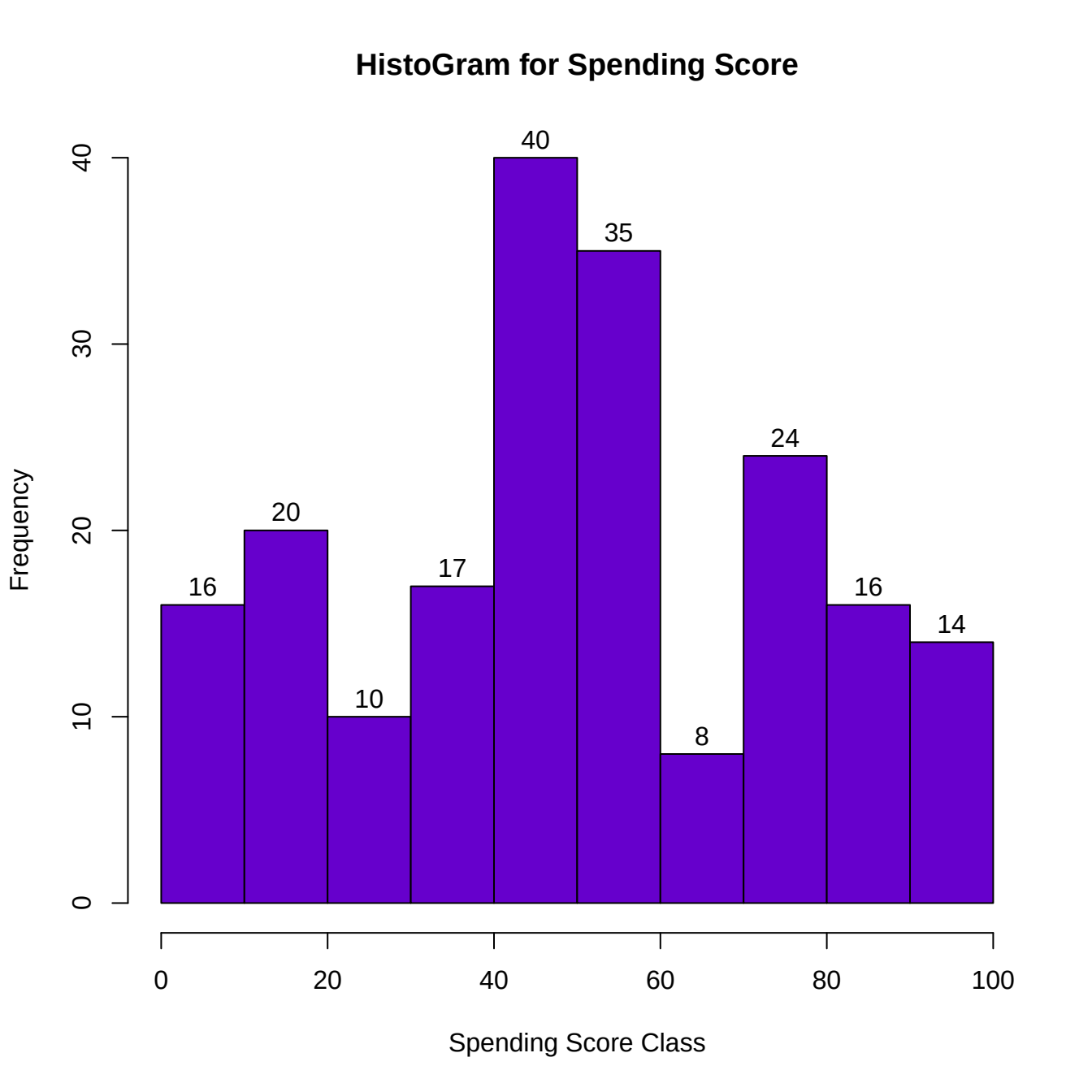

Histogram to Show Count of Age Class
Frequency
0
5
10
15
20
25
30
35
17
21
24
36
24
12
26
11
12
5
12
20
30
40
50
60
70
Age Class

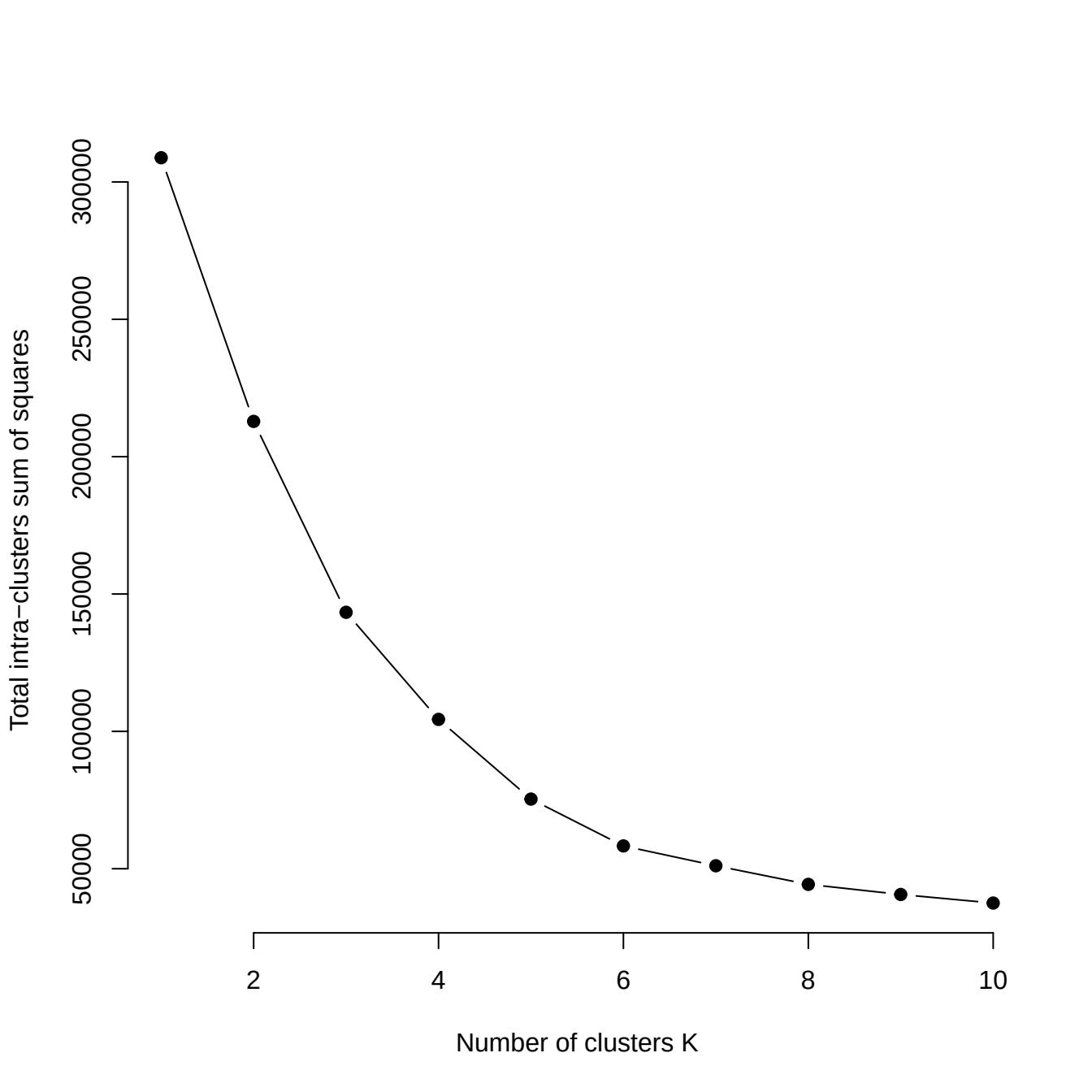

Boxplot for Descriptive Analysis of Age
70
60
50
40
30
20

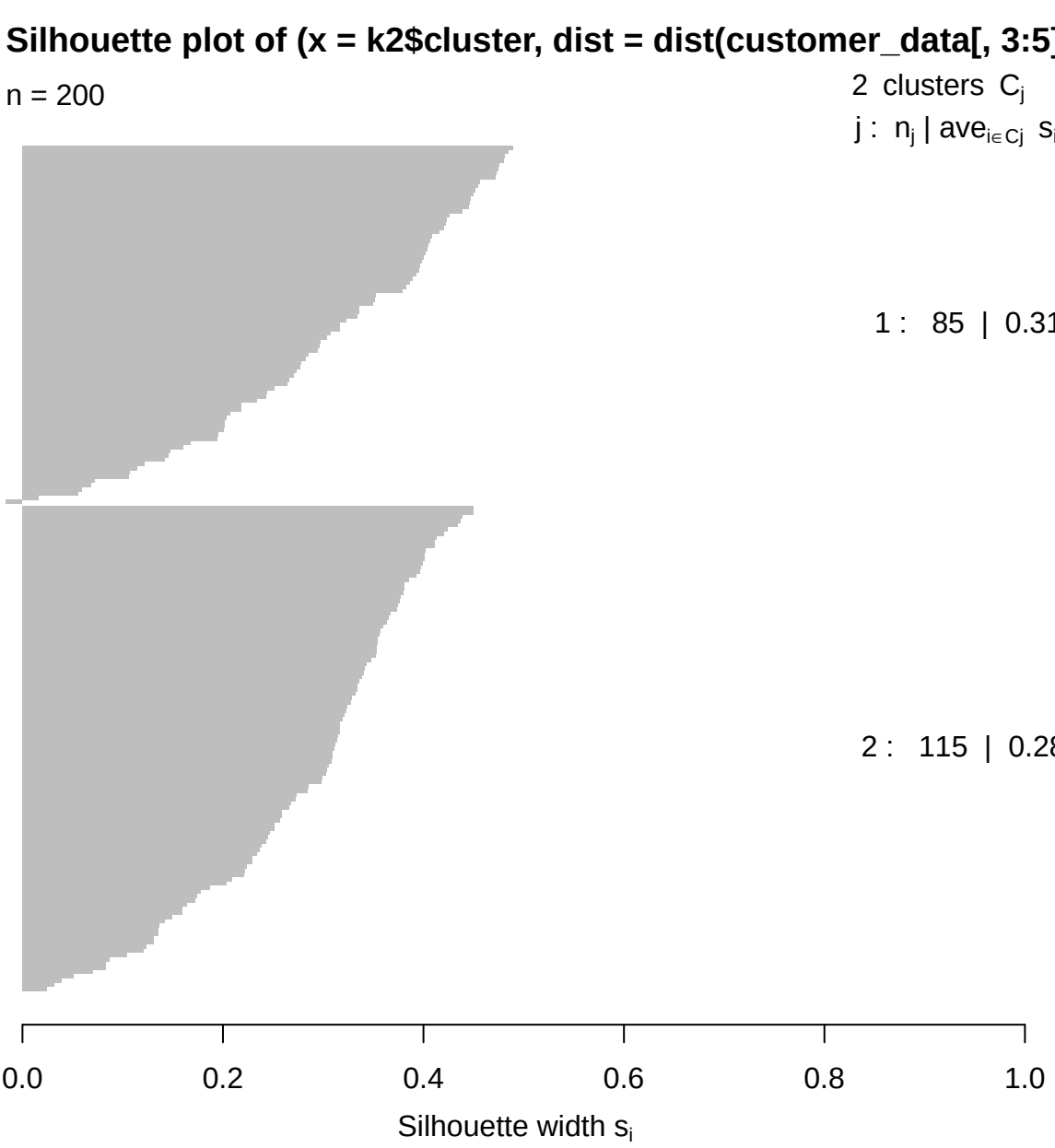

Histogram for Annual Income
Frequency
0
5
10
15
20
25
30
35
16
16
18
24
24
28
36
16
8
6
4
2
2
20
40
60
80
100
120
140
Annual Income Class

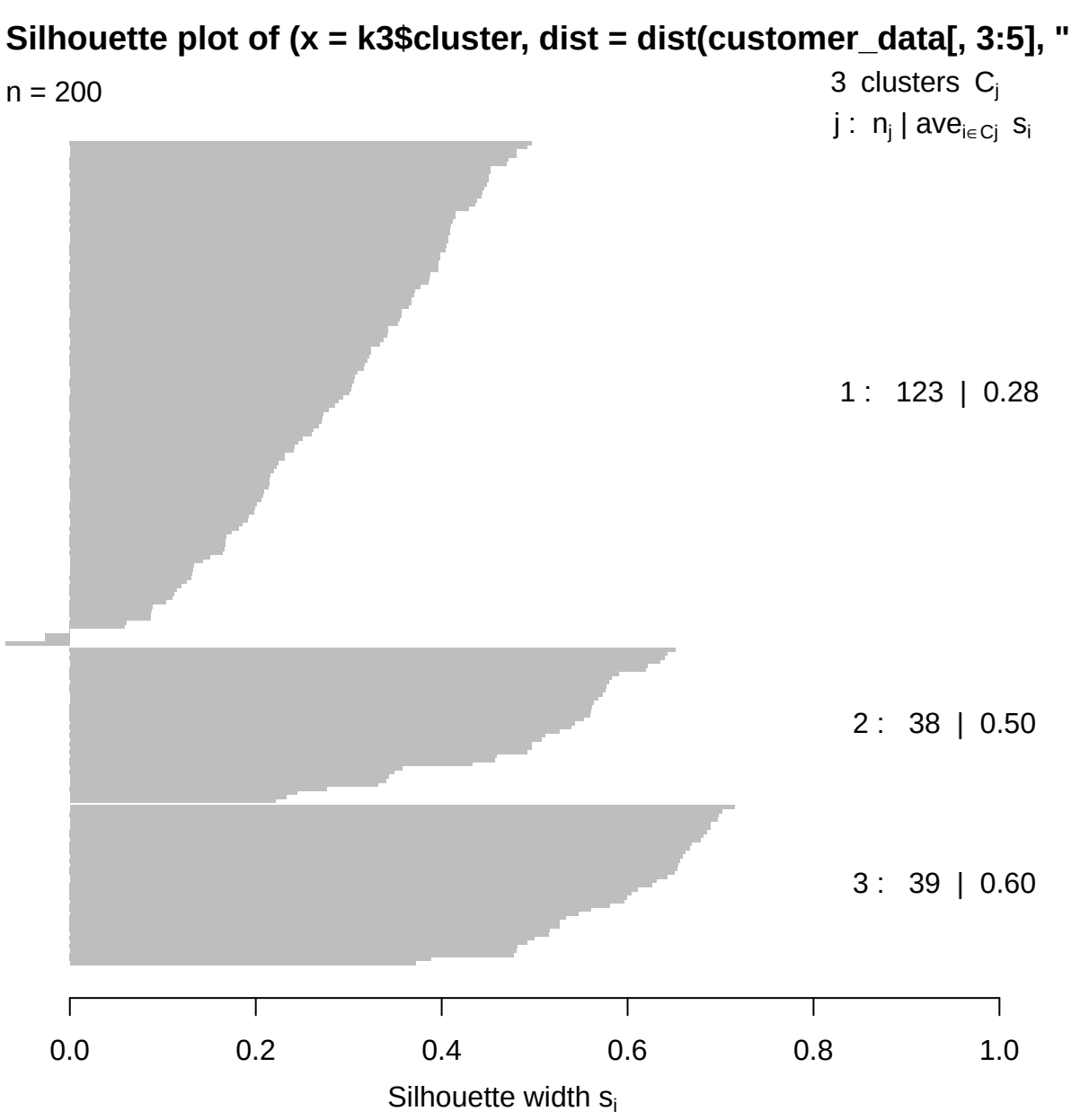

Density Plot for Annual Income
Density
0.000
0.005
0.010
0.015
0
50
100
150
Annual Income Class


BoxPlot for Descriptive Analysis of Spending Score
0
20
40
60
80
100

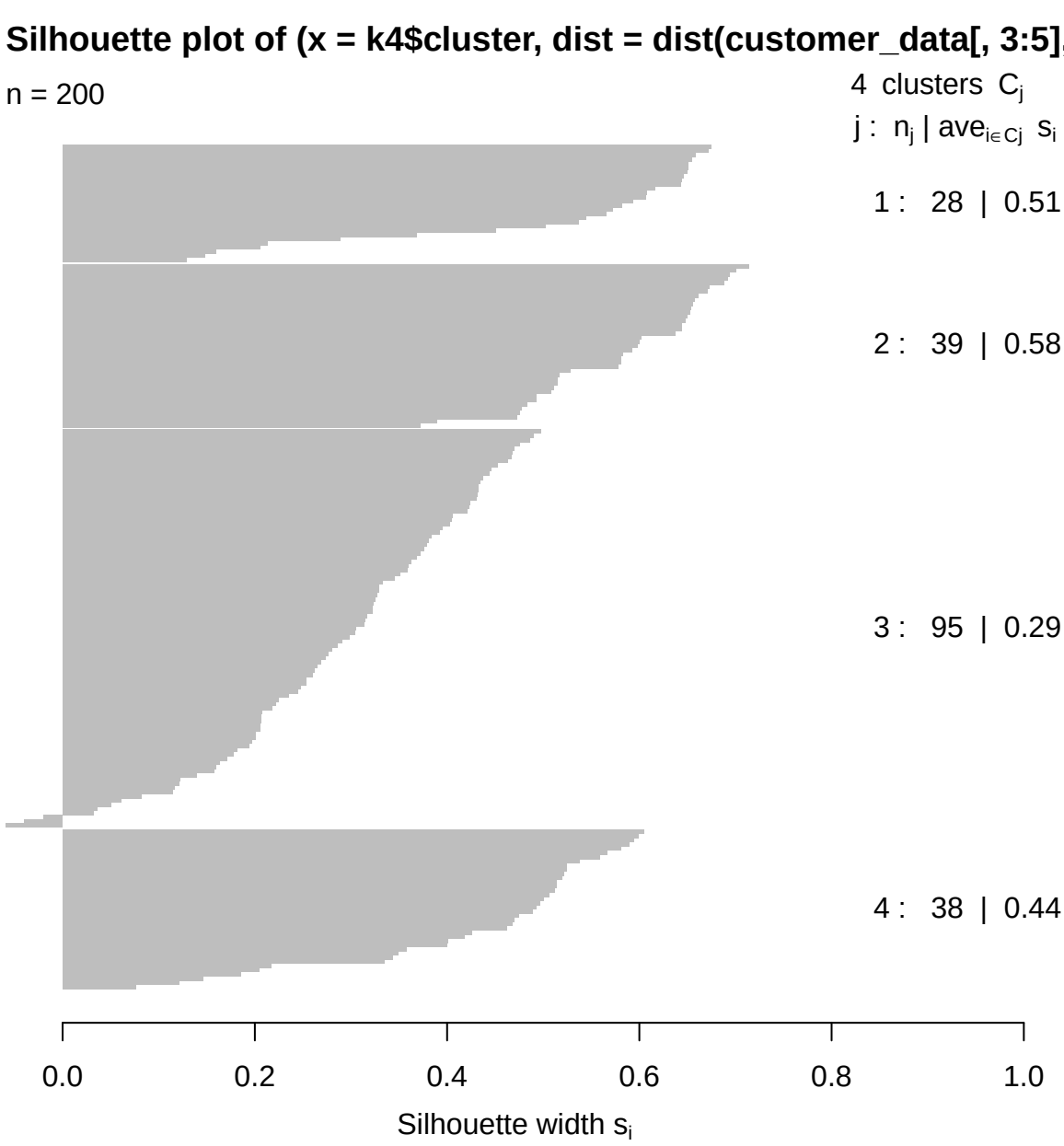

HistoGram for Spending Score
Frequency
0
10
20
30
40
16
20
10
17
40
35
8
24
16
14
0
20
40
60
80
100
Spending Score Class

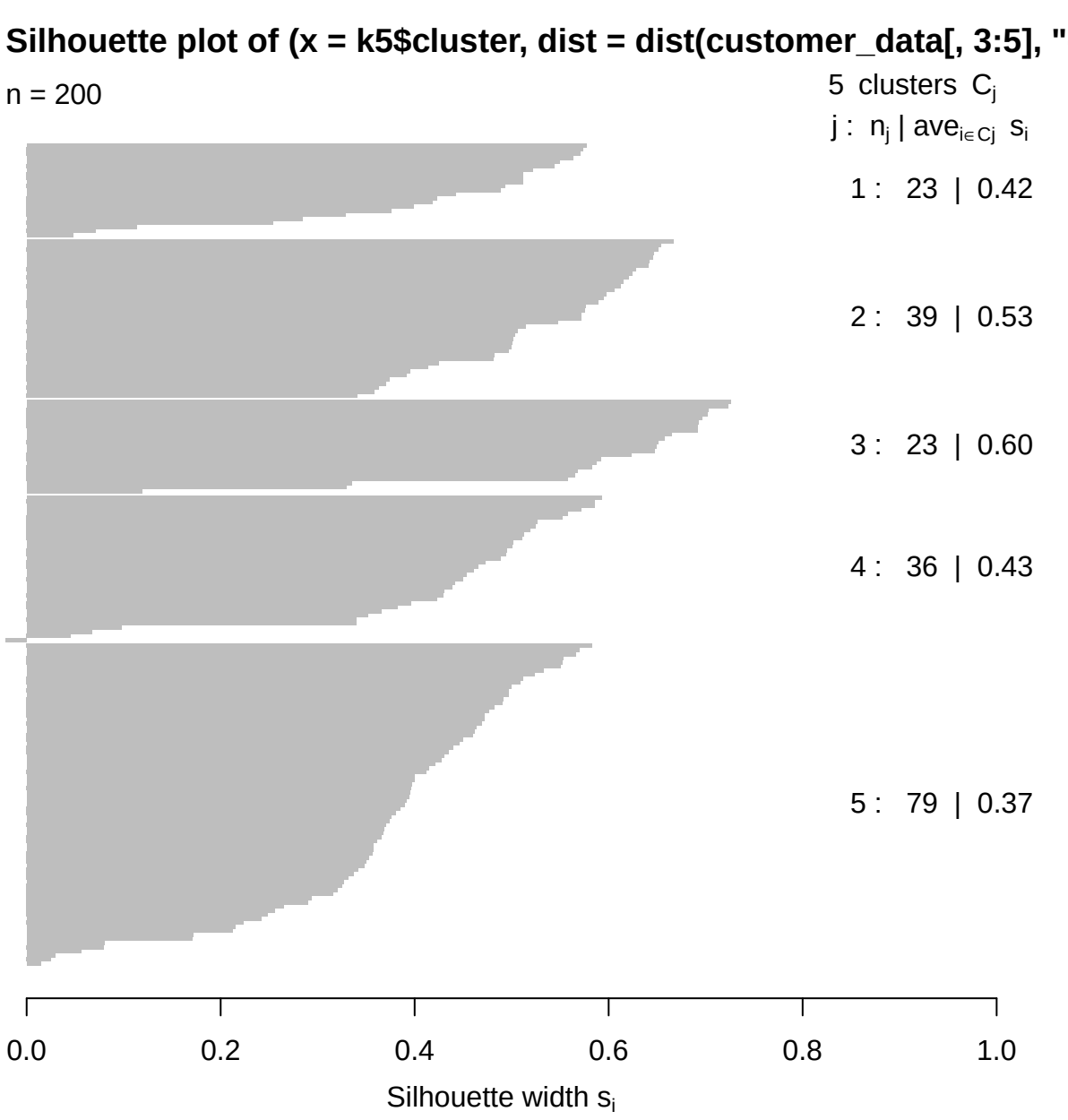

Total intra−clusters sum of squares
50000
100000
150000
200000
250000
300000
2
4
6
8
10
Number of clusters K


Silhouette plot of (x = k2$cluster, dist = dist(customer_data[, 3:5], "
n = 200
2 clusters C_j
j : n_j | ave$_{i \in Cj}$ s_i
1 : 85 | 0.31
2 : 115 | 0.28
0.0
0.2
0.4
0.6
0.8
1.0
Silhouette width s_i
Average silhouette width : 0.29

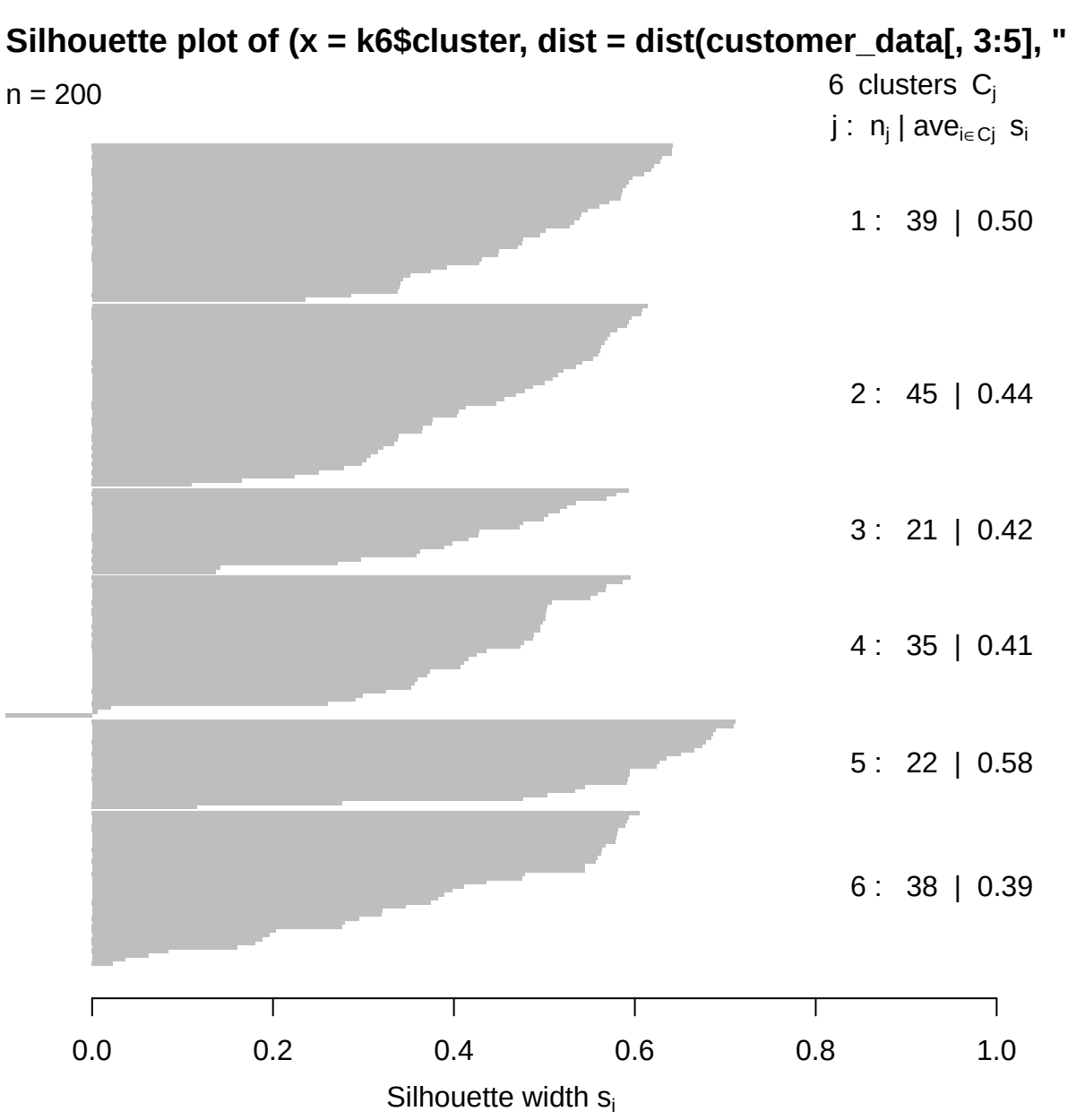


Average silhouette width : 0.38

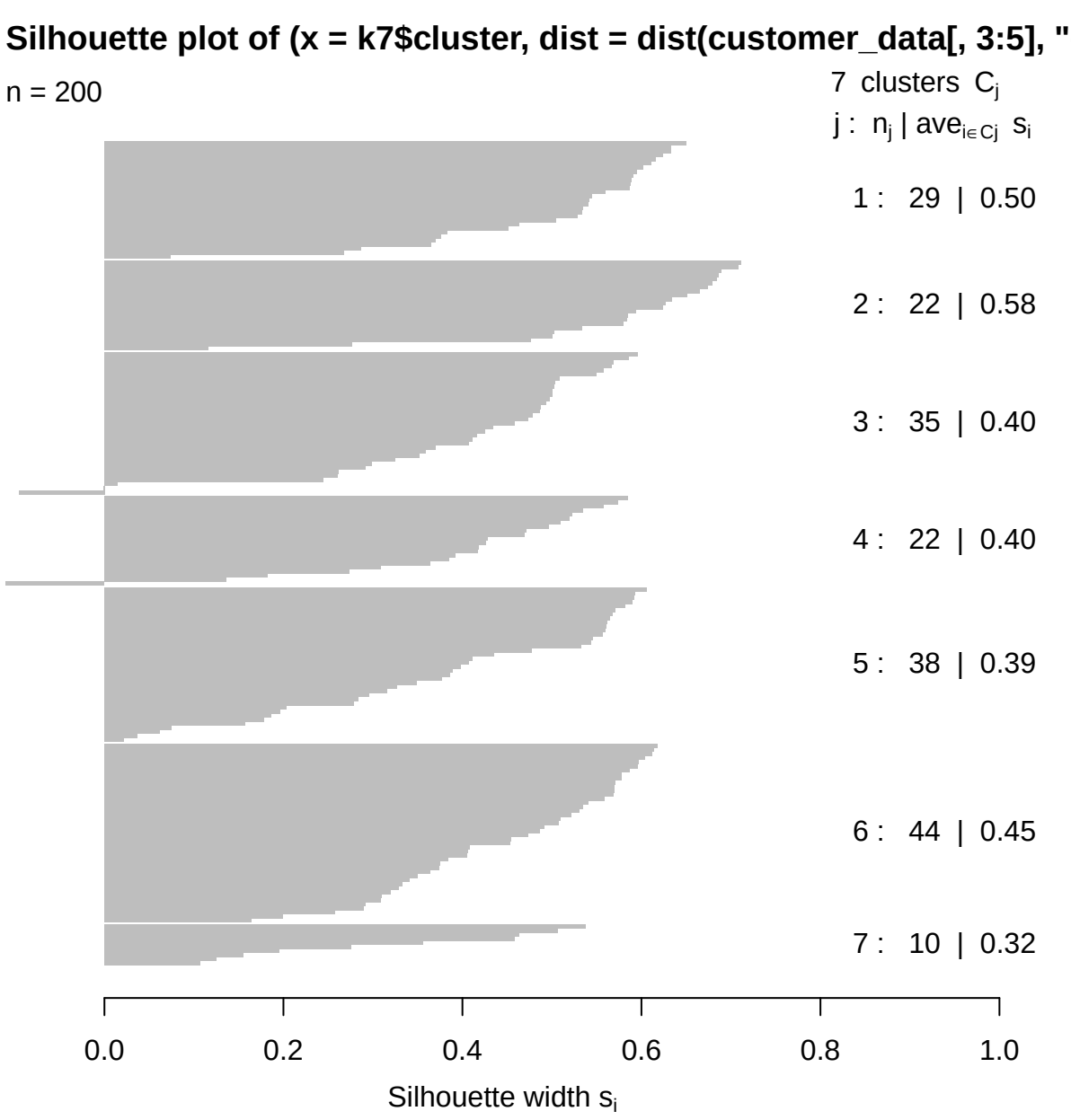

Silhouette plot of (x = k4$cluster, dist = dist(customer_data[, 3:5], "
n = 200
4 clusters C_j
j : n_j | $ave_{i \in Cj}$ s_i
1 : 28 | 0.51
2 : 39 | 0.58
3 : 95 | 0.29
4 : 38 | 0.44
0.0
0.2
0.4
0.6
0.8
1.0
Silhouette width s_i
Average silhouette width : 0.41

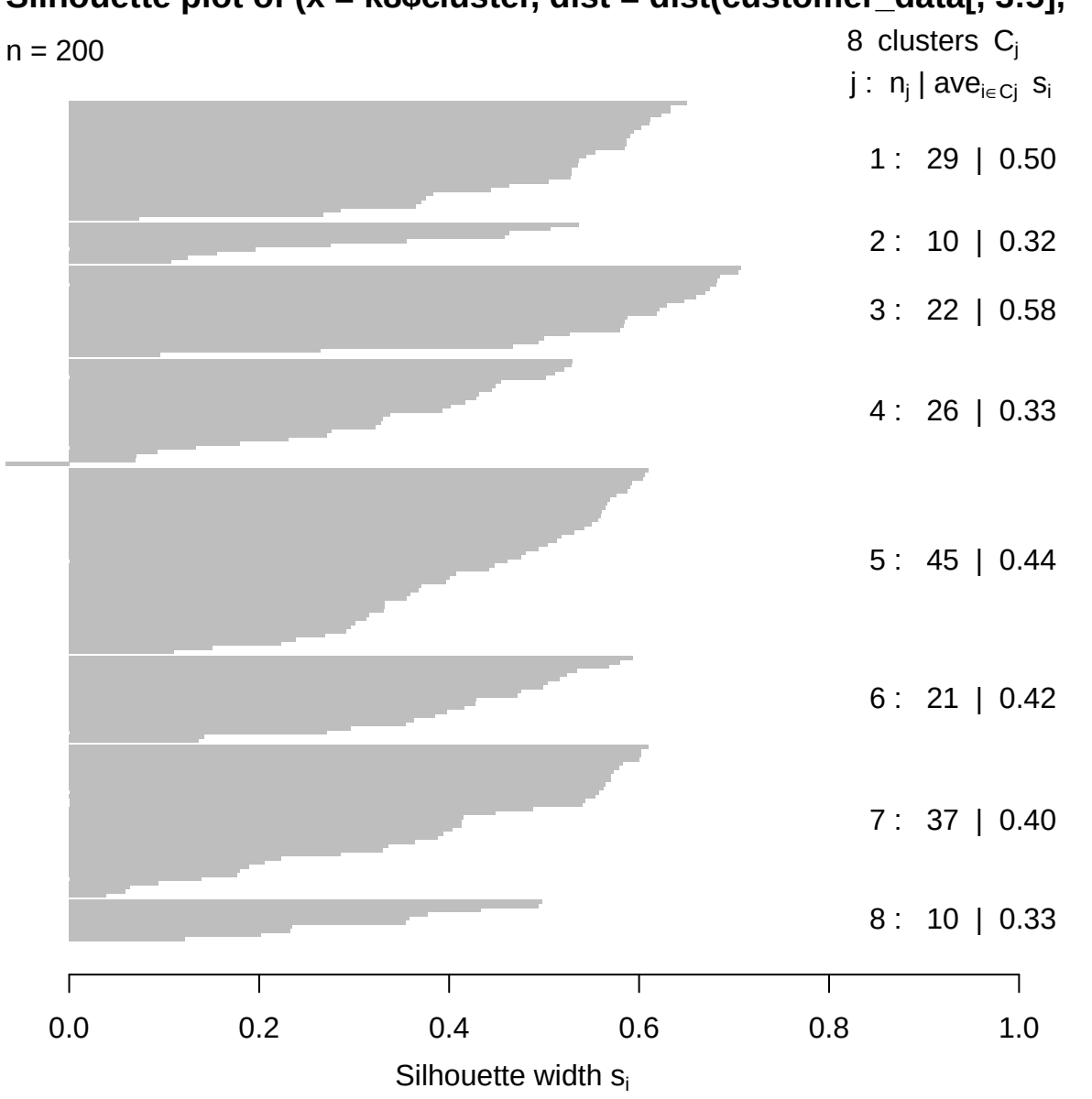


Average silhouette width : 0.44

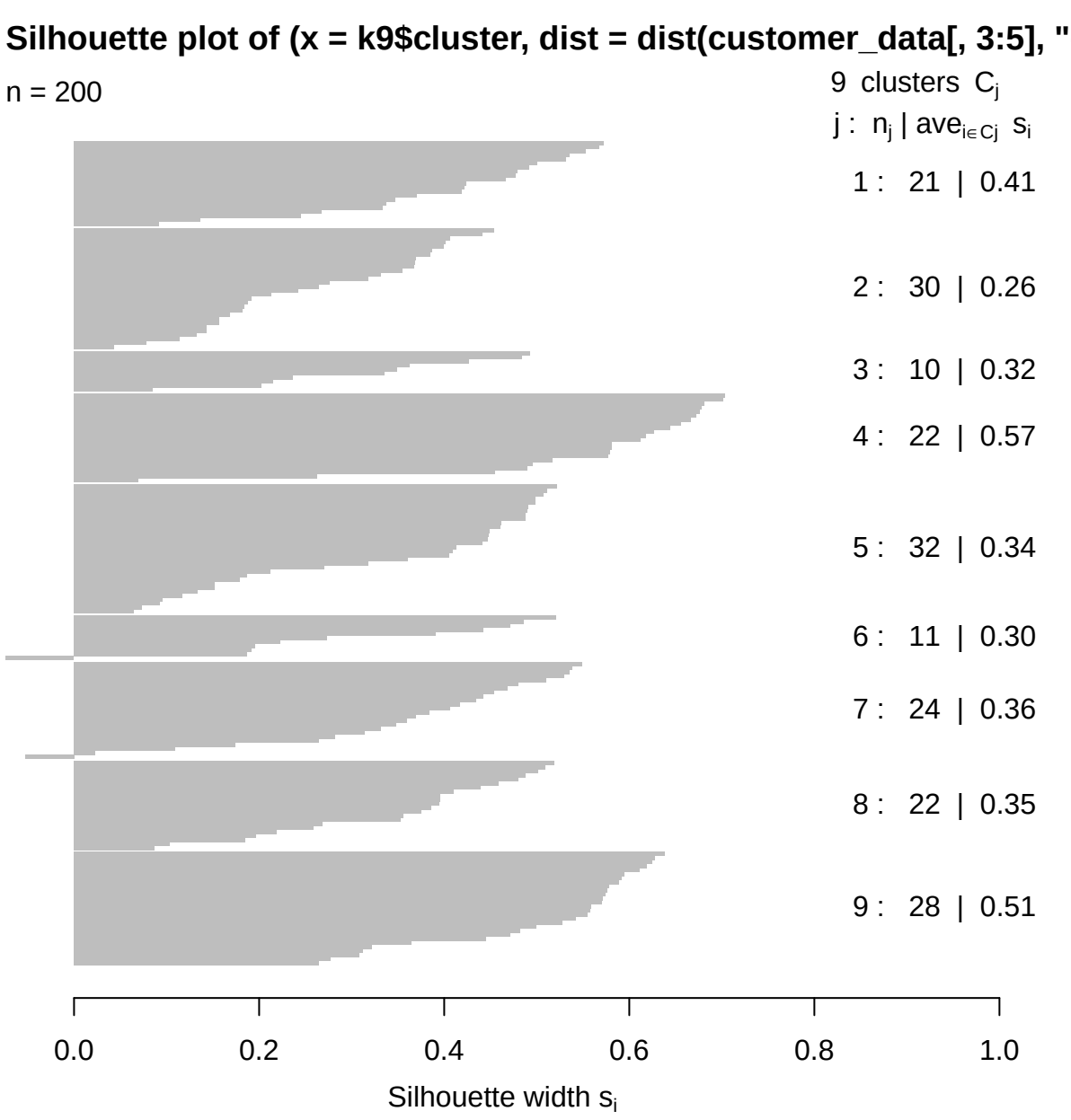


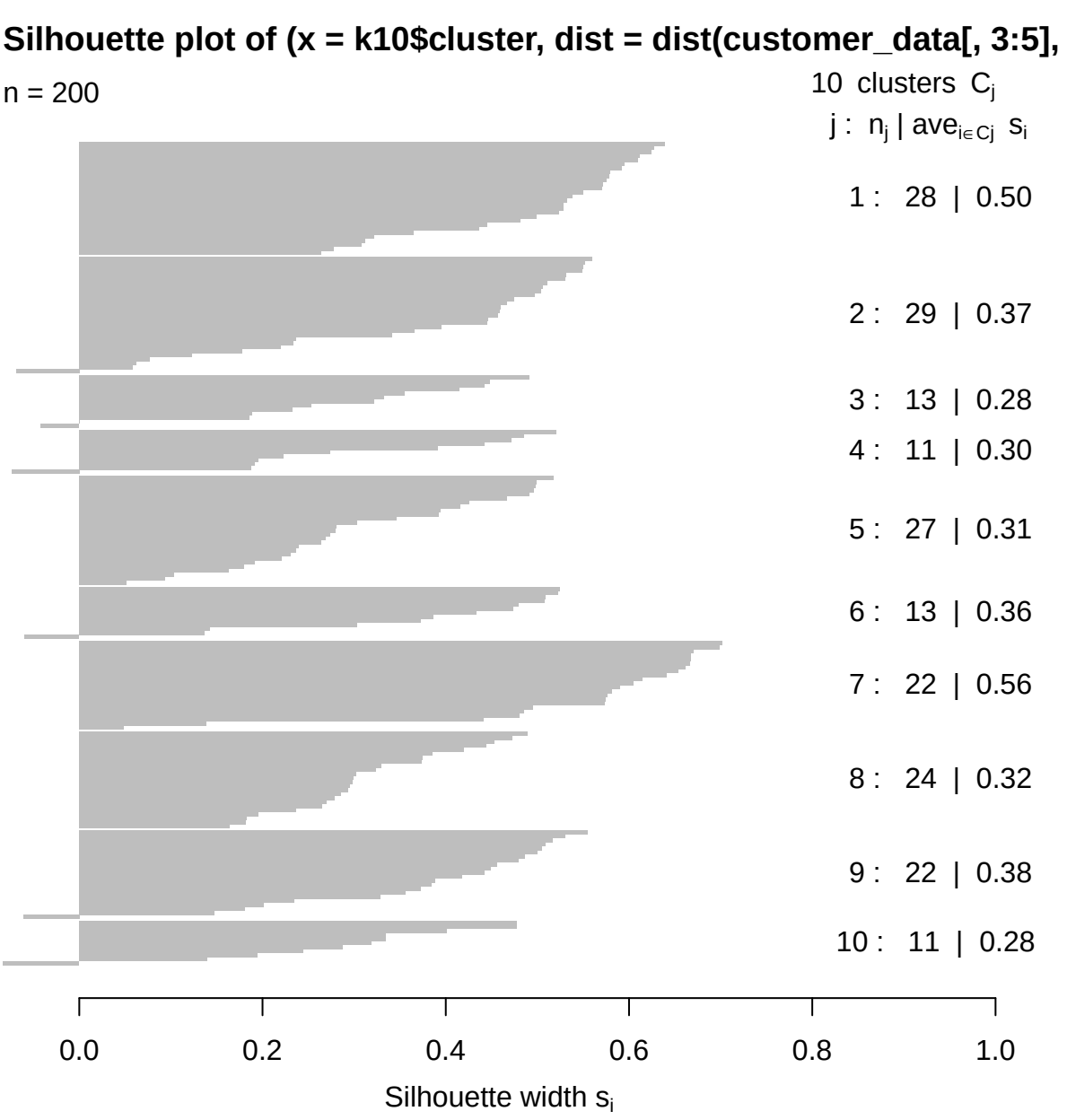
Average silhouette width : 0.45


Silhouette plot of (x = k7$cluster, dist = dist(customer_data[, 3:5], "
n = 200
7 clusters C_j
j : n_j | ave$_{i \in Cj}$ s_i
1 : 29 | 0.50
2 : 22 | 0.58
3 : 35 | 0.40
4 : 22 | 0.40
5 : 38 | 0.39
6 : 44 | 0.45
7 : 10 | 0.32
0.0
0.2
0.4
0.6
0.8
1.0
Silhouette width s_i
Average silhouette width : 0.44

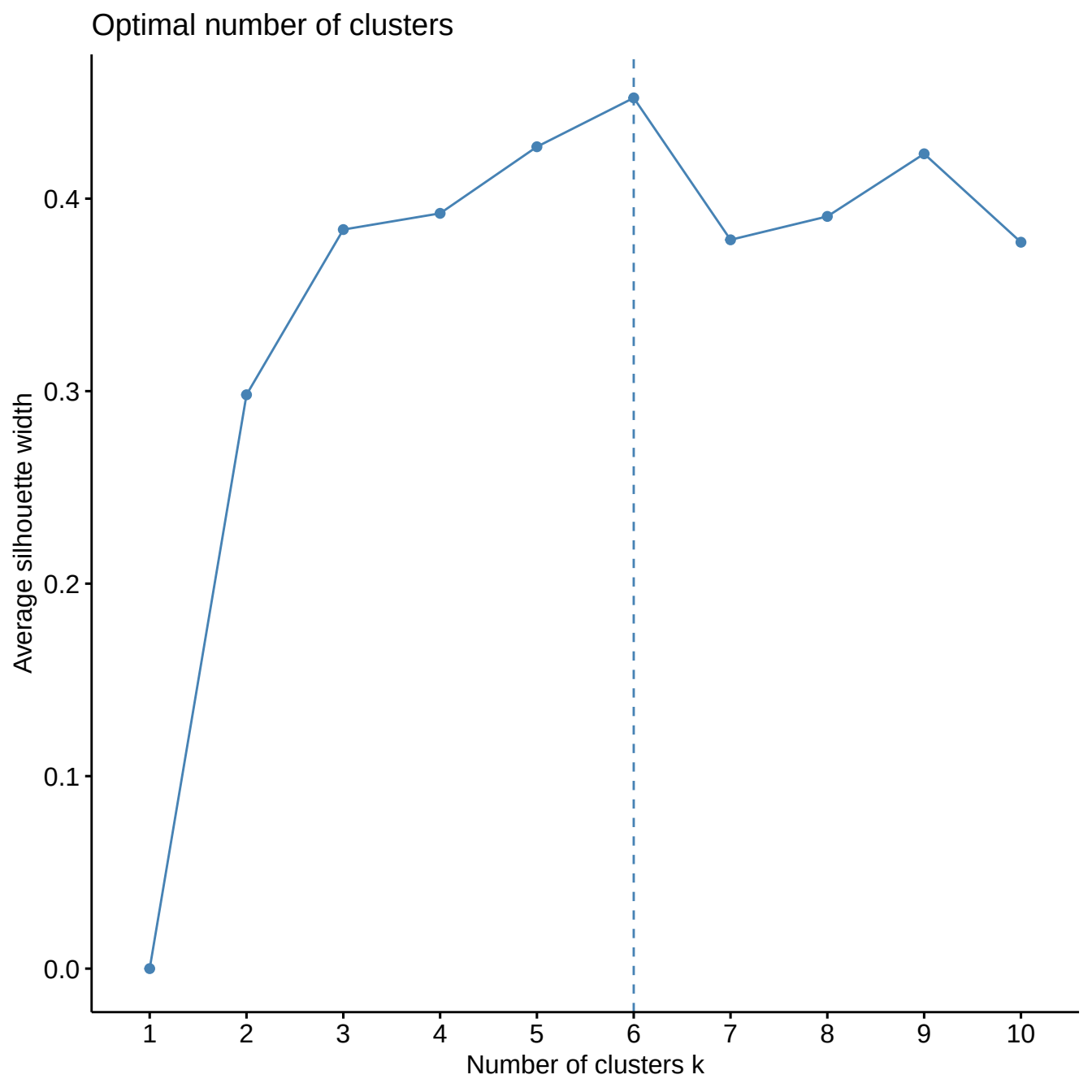


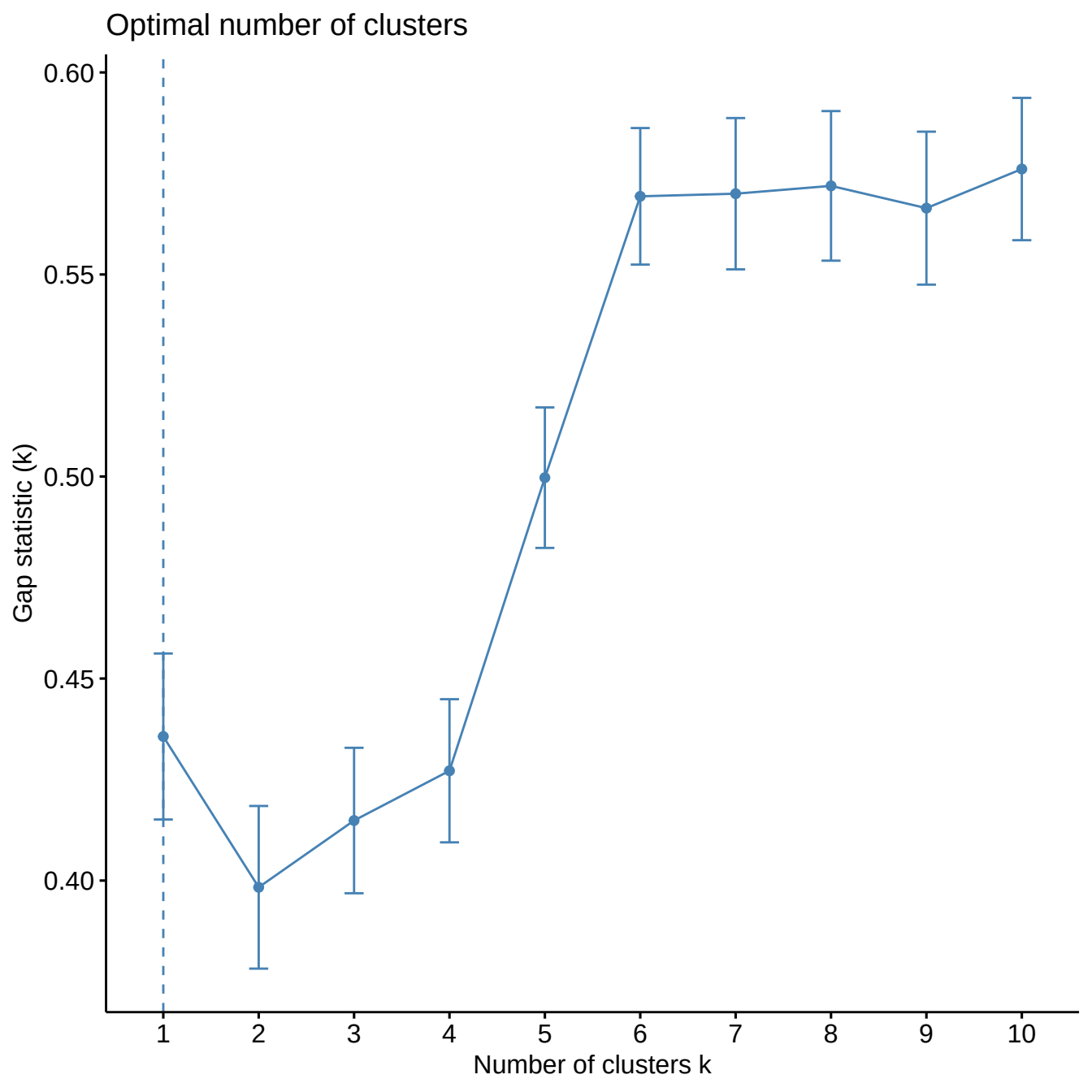
Average silhouette width : 0.43

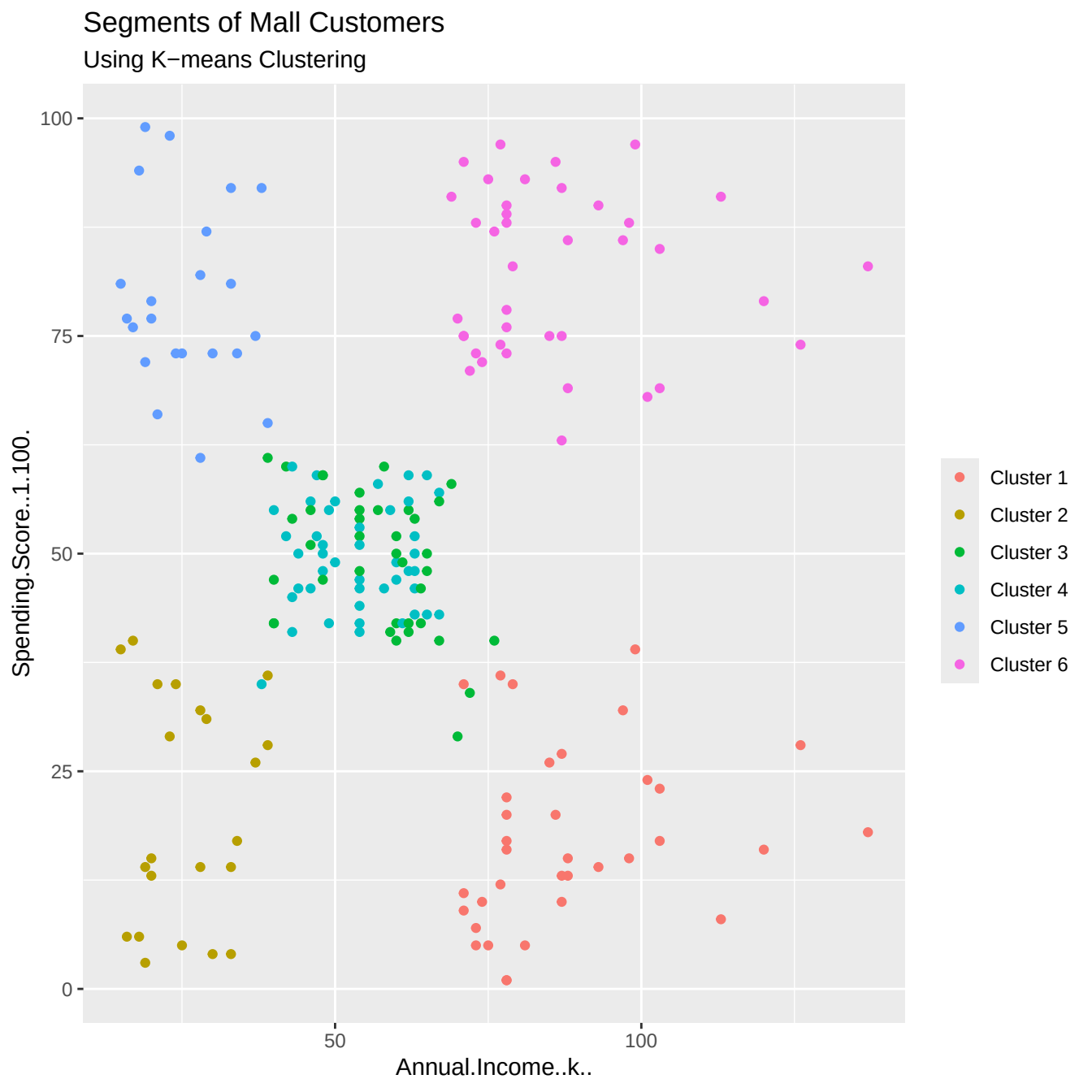

Silhouette plot of (x = k9$cluster, dist = dist(customer_data[, 3:5], "
n = 200
9 clusters C_j
j : n_j | $ave_{i \in Cj}$ s_i
1 : 21 | 0.41
2 : 30 | 0.26
3 : 10 | 0.32
4 : 22 | 0.57
5 : 32 | 0.34
6 : 11 | 0.30
7 : 24 | 0.36
8 : 22 | 0.35
9 : 28 | 0.51
0.0
0.2
0.4
0.6
0.8
1.0
Silhouette width s_i
Average silhouette width : 0.39

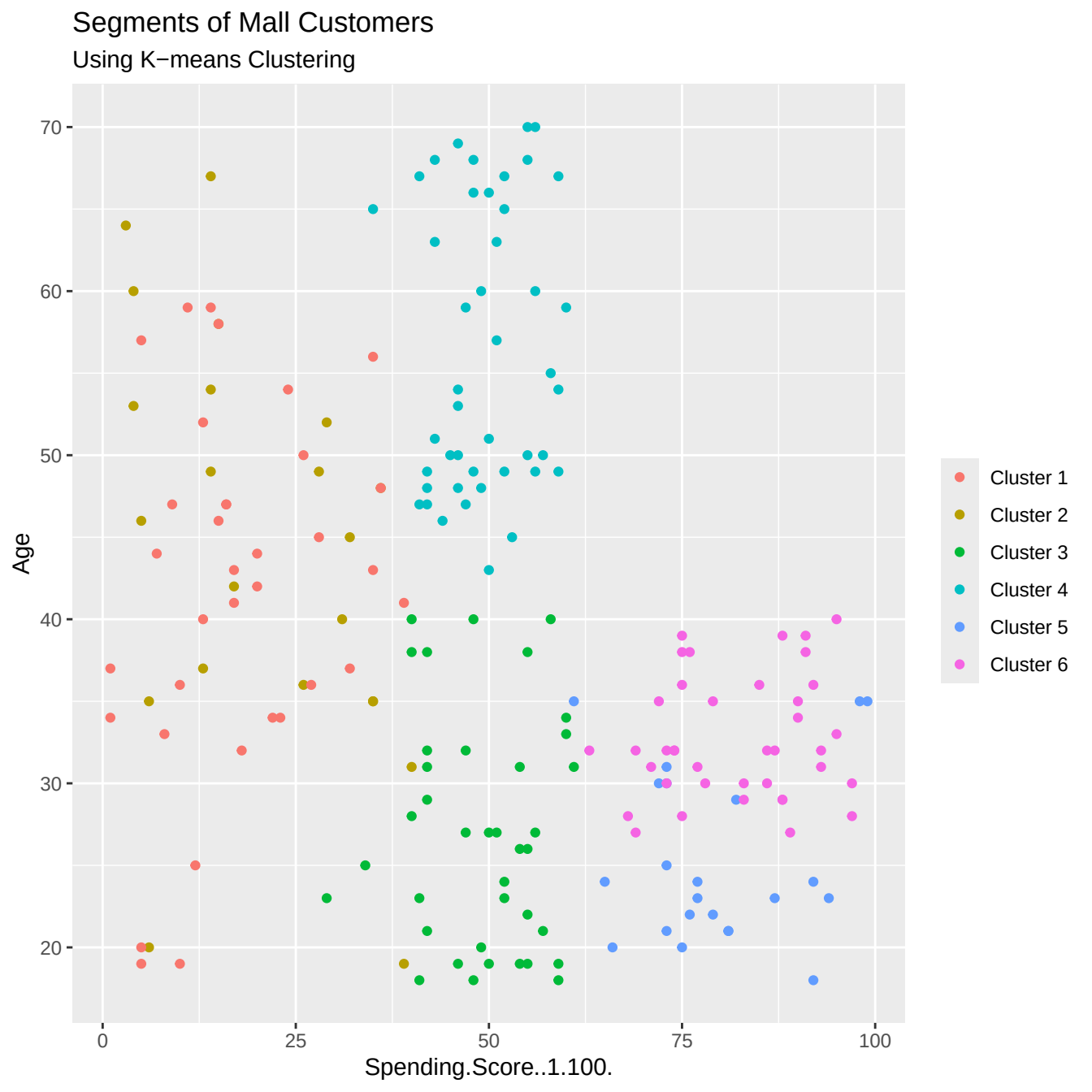


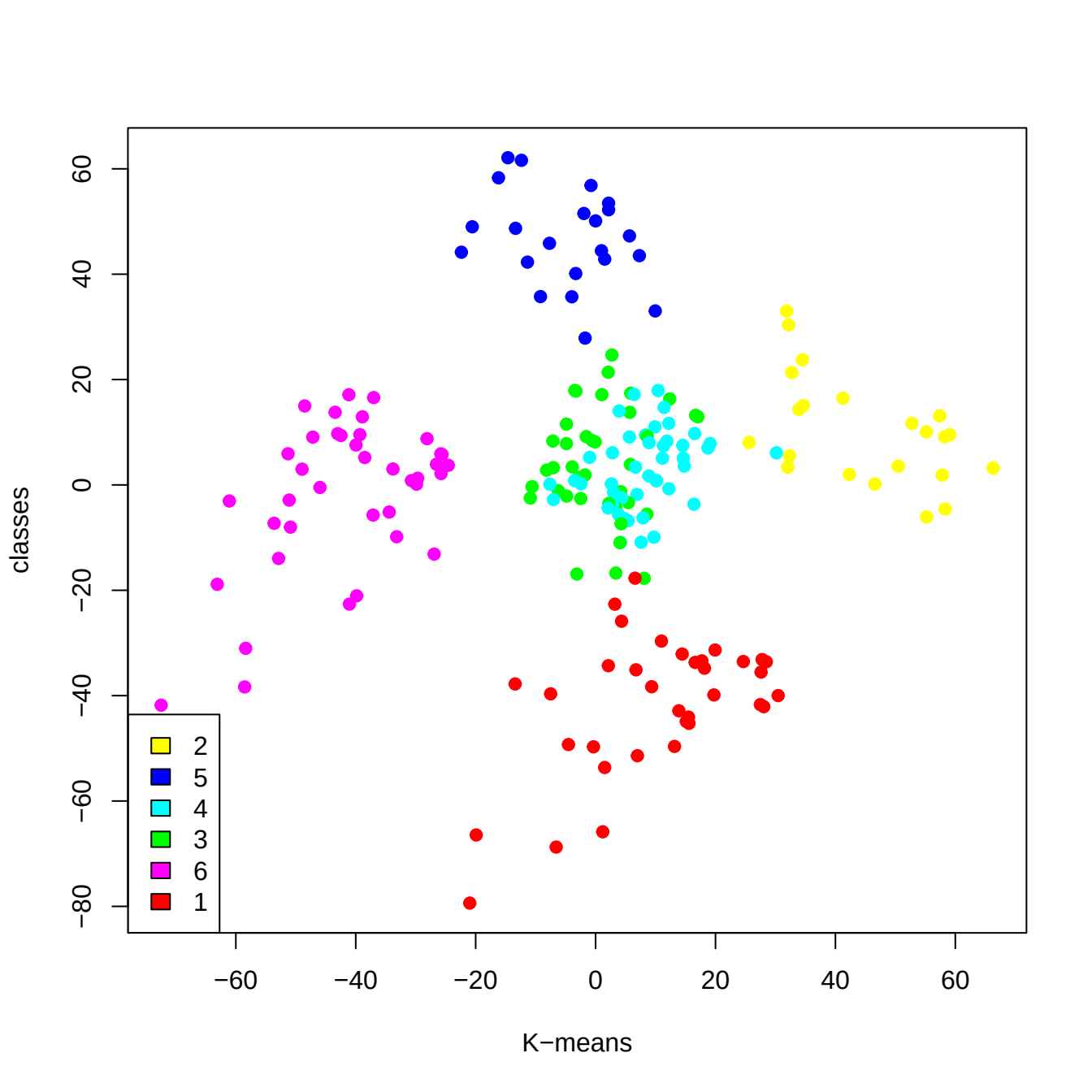
Average silhouette width : 0.38


Optimal number of clusters
Average silhouette width
0.0
0.1
0.2
0.3
0.4
1
2
3
4
5
6
7
8
9
10
Number of clusters k


Optimal number of clusters
Gap statistic (k)
0.60
0.55
0.50
0.45
0.40
1
2
3
4
5
6
7
8
9
10
Number of clusters k


Segments of Mall Customers
Using K−means Clustering
Spending.Score..1.100.
Annual.Income..k..
100
75
50
25
0
50
100
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6


Segments of Mall Customers
Using K−means Clustering
Age
70
60
50
40
30
20
0
25
50
75
100
Spending.Score..1.100.
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6


classes
K−means
−80
−60
−40
−20
0
20
40
60
2
5
4
3
6
1