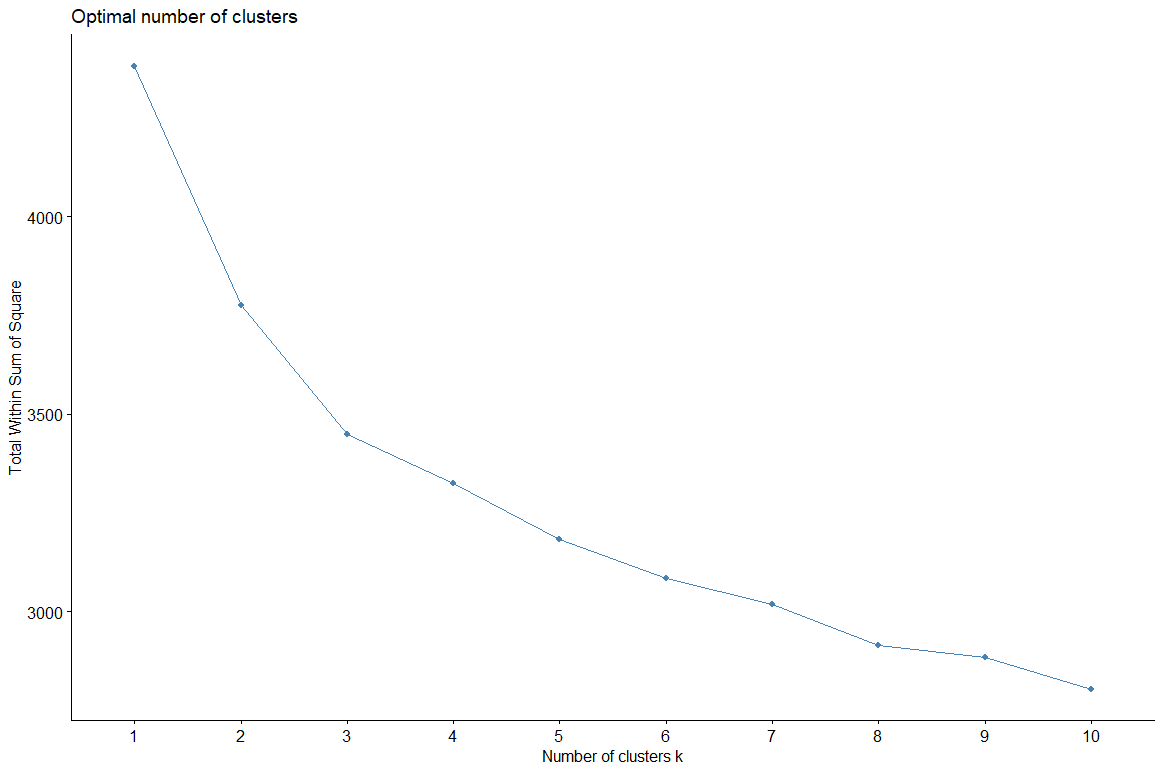


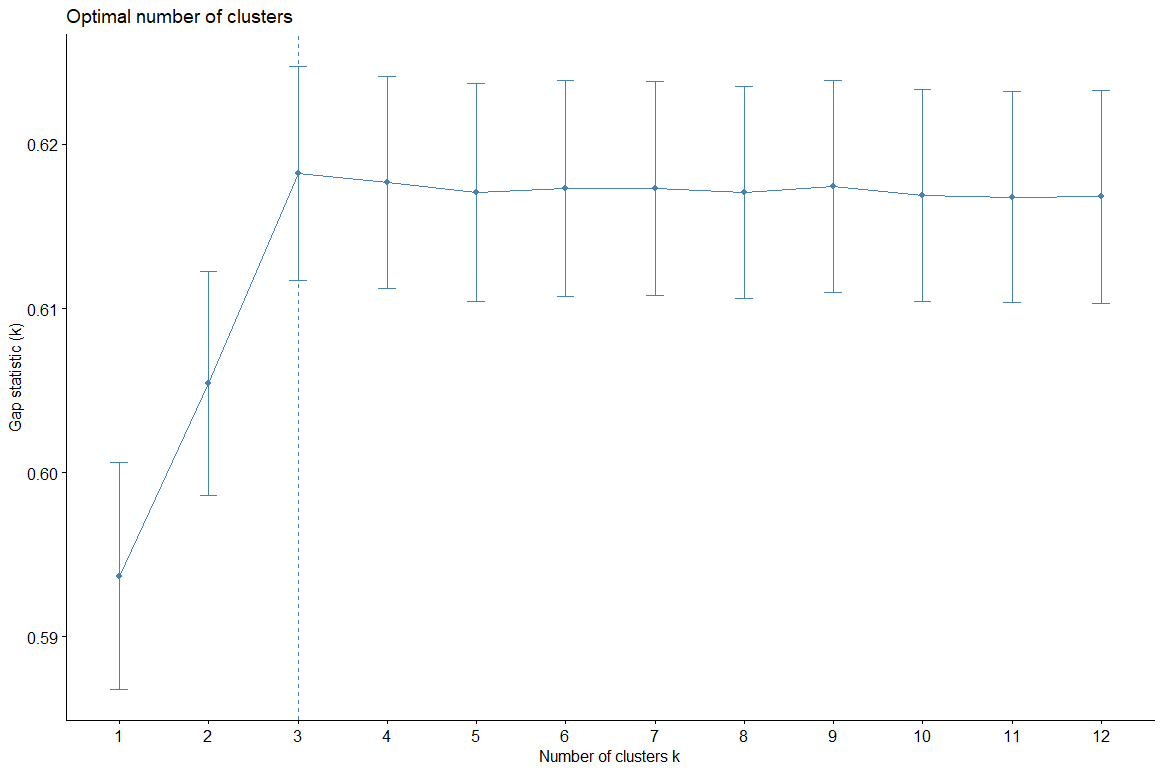
Group 1 - “Family-oriented”  
Group 2 - “Social butterflies”  
Group 3 - “Neutral”

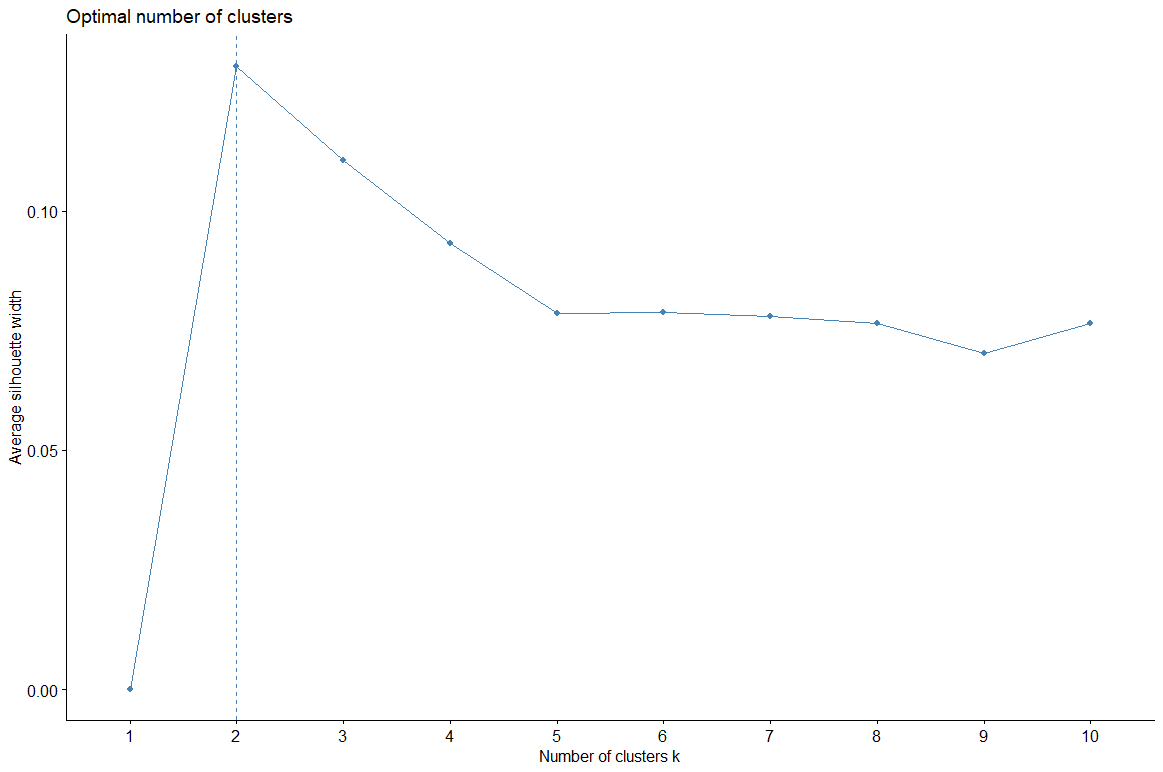
Group 4 - “Anti-social”

The code and graph above show the clustering of the data into four groups, the names we have decided on are above. The first group has a focus on family-oriented social gatherings, since they are the highest rated on visiting relatives and out with the family variables. The second group has a focus on other social events and settings and rated high on live music shows, go out with friends, and go to bars variables. The third group was a middle ground for most of the variables, therefore making them neutral and not leaning to much to any of the other groups. The last group had very low ratings for the variables for social settings, therefore seemed to be a more anti-social group. The graph also shows the separation of these groups, with the third, neutral, group being in the middle of the other groups.

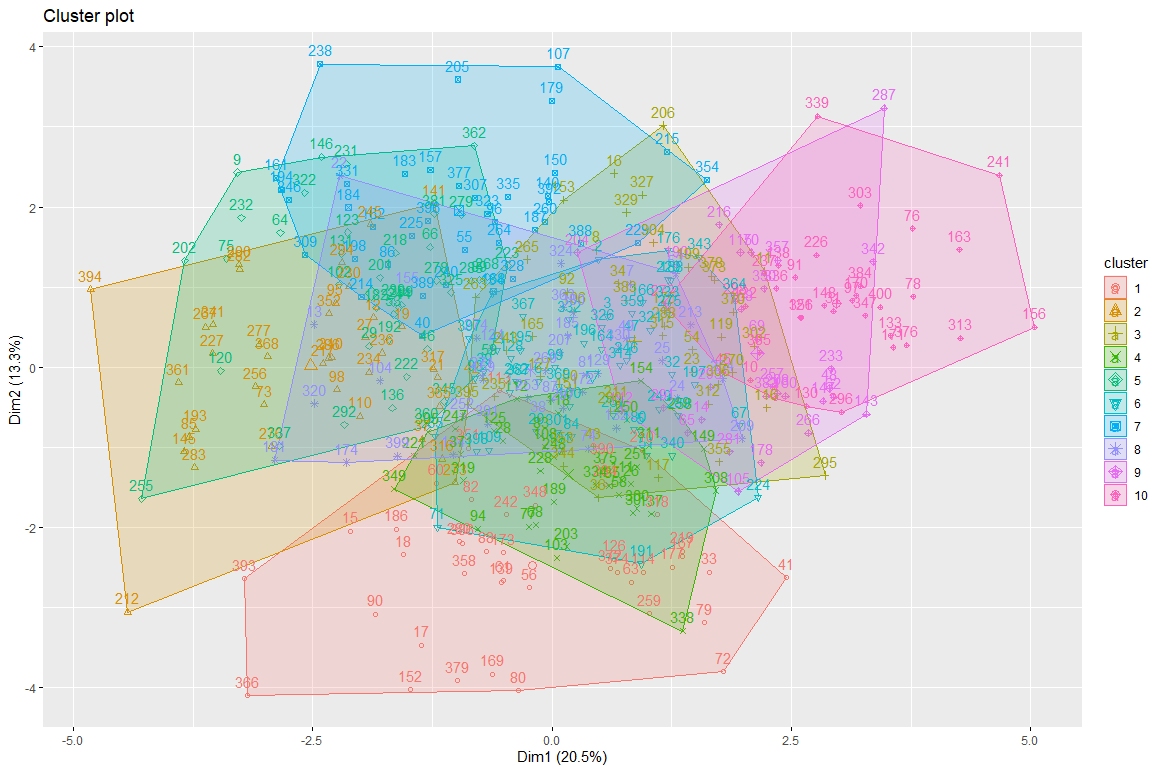
Assuming that we are the Walmart marketing team, for group 4, the anti-social group, we could further push the car pickup system so that those individuals may order ahead and pick up their items without having to go inside and lead to interactions, and the advertisements could be made through radio and TV since this group has high usage for both of those forms of media. For group 1, the family-oriented group, we could create advertisements that depict a family setting or family social functions and how Walmart can provide the items to make these events happen, such as food, tables, lawn chairs, family-oriented games or board games, etc, and would be shown on TV, newspapers, and the radio since this group uses these mediums often. For group 2, the social butterflies group, we could create advertisements depicting social settings such as supplies for super bowl parties, drinking with friends, and traveling supplies, and the advertisements could be shown through the radio and in magazines. All of these advertisements would reach the different sides of group 3, the neutral group, offering something for everyone within that group. However, this group may not be necessary to have…

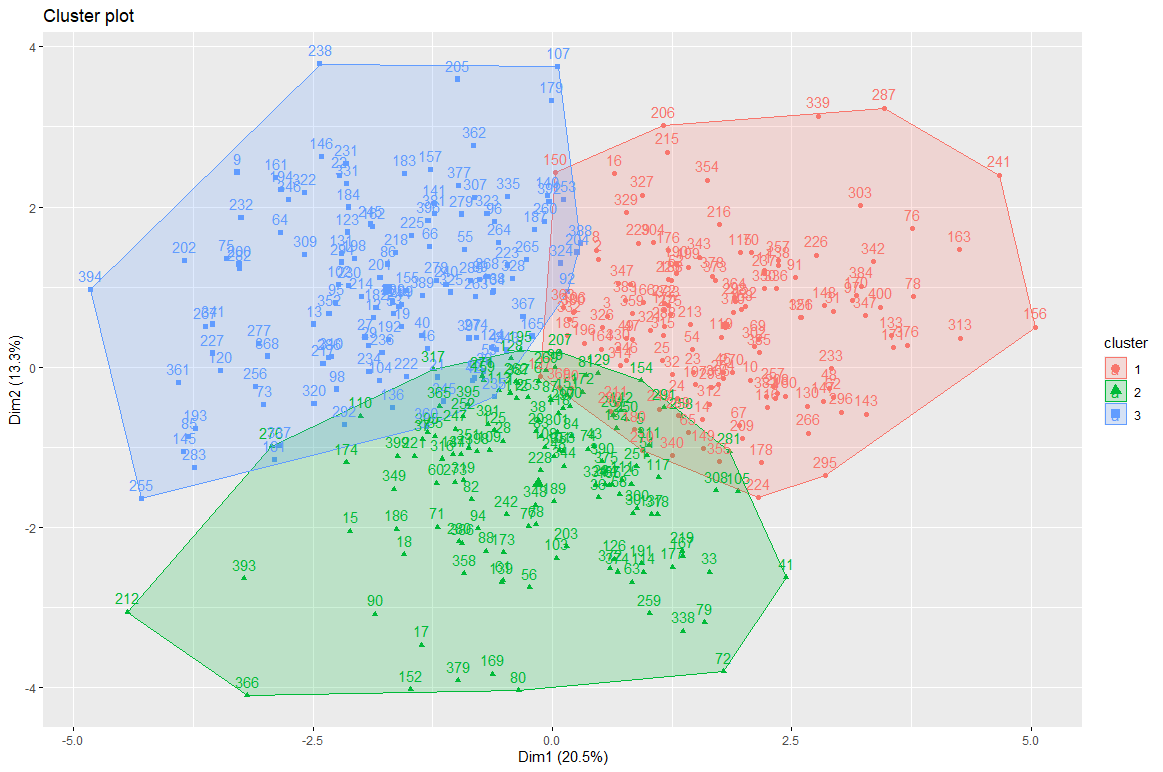


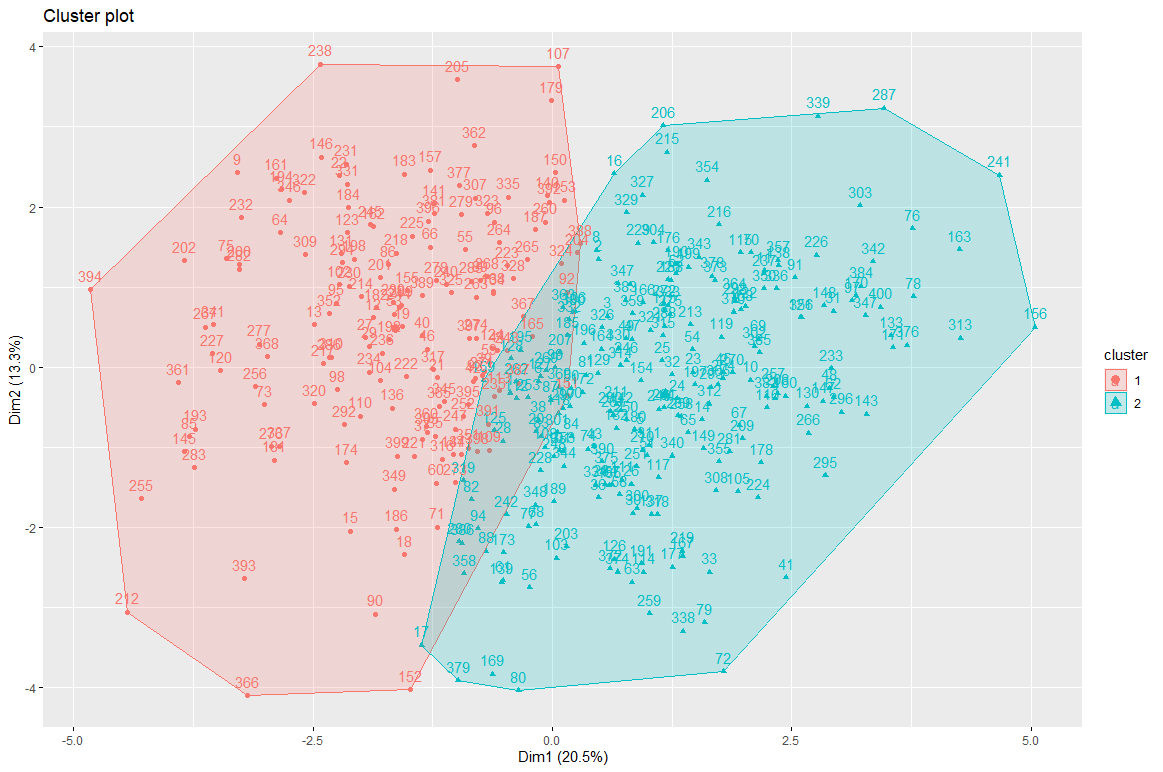




Using silhouette, gap, and elbow method to determine the optimal number of groups, we get 10, 3, and 2 as the optimal number of groups to have for clustering this data. We can graph these to determine which number of clusters creates a better graph. In the graphs, 10 clusters appears to be too many and the groups have too many overlapping features for the groups to be unique enough to identify. 3 clusters shows three strong groups, which follows with the observations made from the initial grouping and data analysis above. 2 clusters is too broad for the prompt of assuming that we are Walmart trying to target our store to customers. From above, 4 groups results in one group having too many overlapping values with the other 3 groups, therefore 3 seems to be the optimal number of clusters for grouping the customers in the data.







R Studio code  
