I have labelled the dataset in R using the file provide in Moodle but, I used numbers to represent them in the list because, some labels have the same name and it becomes confusing when fusing clusters. So doing this makes it easier to visualise the results.

I have written but commented out the code for the list to have names, which can be uncommented if you want to check but it is more confusing to read.

3. The single-linkage algorithm produced more pairs of clusters at then start of thee algorithm where you can see the first 32 iterations are just pairs of clusters joining to make a larger cluster. Then after this each cluster is added one at a time to the cluster [1,2,3,9] formed at iteration 33.
   This when compared to the average and complete linkage algorithms is completely different. The average and complete linkage algorithms join clusters together till there are generally around 4 to 5 clusters making up a single cluster e.g. [24,29,30,54]. This makes the dendrogram seem more balanced when compared to the single-linkage algorithm.
   The centroid-linkage algorithm acted slightly differently from the other algorithms first pairing cluster together but then adding clusters in a pyramid format before joining the final clusters together (See image below).

CDLM	list [63]	List of length 63
[[1]]	integer [2]	50 51
[[2]]	integer [3]	49 50 51
[[3]]	integer [2]	57 58
[[4]]	integer [2]	21 22
[[5]]	integer [2]	35 36
[[6]]	integer [3]	35 36 37
[[7]]	integer [2]	1 2
[[8]]	integer [2]	61 62
[[9]]	integer [3]	60 61 62
[[10]]	integer [4]	60 61 62 64
[[11]]	integer [5]	59 60 61 62 64
[[12]]	integer [6]	59 60 61 62 63 64
[[13]]	integer [2]	12 13
[[14]]	integer [3]	12 13 14
[[15]]	integer [4]	11 12 13 14
[[16]]	integer [5]	11 12 13 14 15
[[17]]	integer [6]	11 12 13 14 15 16
[[18]]	integer [7]	11 12 13 14 15 16 ...
[[19]]	integer [8]	11 12 13 14 15 16

CDLM[[18]]

4. I used the K values of 2,3,4,5 and 6. When K is 4 the reduction in variation starts reducing at a lower rate, using the elbow method, suggesting that this is the K that should be used for clustering. However, when the value is 2 there is still a large drop in variation till K = 4 but after this point it reduces slowly.
   Using K = 4 there is a lot of overlapping of data which could be a result of hard assignment or how close the labels in the dataset are linked.

5. In this case, I would say that hierarchical clustering is easier to read in the sense that if you were to cut a dendrogram you could clearly see the separation in clusters. Whereas, the K-means points overlap when plotted although it is an accurate representation of the dataset.

6. In K-means you can choose the number of clusters K by testing different methods of K and looking at which one performs better. You could also find the K at which point the reduction in variation stops dramatically decreasing also known as the elbow method.

   In hierarchical clustering you don't choose the amount of clusters, you just pair clusters that are most dissimilar according to a given linkage algorithm (either agglomerative or divisive). So I am not sure how you would set the amount of clusters as the question asks using this particular algorithm.
   Unless you decide at which point in the dendrogram (or the lists in the case of my own algorithms) you choose to cut a line through it providing you with a set amount of clusters depending on where the cut were to occur.