Regarding classes, we followed our development plan almost exactly. We had a user interface class, one to web scrape the spotify charts we wanted to cluster and one to save the different audio features. We had a separate class to calculate the K-means values and distances between songs, but also to compile the minimum spanning tree and to visualise it.

In visualising the minimum spanning tree, we ran into a problem. As we used many different features to calculate the positions of songs in the minimum spanning tree, visualisation in 2D did not add much to the clustering of songs. Songs of the same colour were placed all around the graph, which did not clarify the positions of songs compared to one another much. However, we tried to solve this problem through 3D visualisation. Looking at the clusters from different angles made them much clearer, even though it still does not cover the many different dimensions over which we calculated the clusters.

To validate our clusters, we used the Silhouette method, from which we only got a very small value. Using different K-values did not change this value significantly. This was due to the Sihouette method’s use of the distance between songs within a cluster. Our clustered songs were not necessarily placed together in space, but coloured alike, which caused the validation method to not work correctly.