The task was such as to check the performance of the different clustering algorithms. kMeans provided with a steady and plain circumstancial base of all the entrances based on the nature of the 210 entries we were provided with. The actual graph provided us with a straight and plain oversight without garnering any factual insight upon the task in progress. However, it did provide us with the most base functionality and allowed us to have an overview while fact-checking the provided text of the Train/Test results. The numbers of the evaluation itself are, by design, in the lower accuracy. This is mainly because of the number of entries and the lack of an overall common factor. The entries, while sharing a general connection based on their different classes, have very varying results. As such, the r2-accuracy is of low accuracy.

This has mainly to do with the use of all 210 entries. A smaller amount of entries may have resulted in higher accuracy (as it assumes higher probability for differentiation between the different results). While taking smaller numbers would have been possible, it was decided against it. These are non-altered results based on what we were provided with.

GMM-methods varied a bit more. The four entries were generally providing a similar answer with one exception.  
In order to not change the input gained from kMeans, all 210 entries were used once more with the same seperation of Test/Train. Again, the entries were graphically shown correctly (or, at least, consistently compared to kMeans and each other) and could be changed for comparison easily.  
However, a problem occurred, and accuracy dropped immensely. While we were able to change the size of the graphical assessment, the learning curve based on the data provided was poor.