

Team assignment.

## Step 2: Clustering

Utilize the same data set from Project 2 Step 1 (SOM) and mine it using clustering. Every team is assigned a clustering algorithm at the end of the document.

Deliver a report and the clustering code or script you have used. The report must contain:

1. The clustering algorithm with all its parameters.  
**5 points**
2. Expectations from clustering – number of clusters expected and actually received, performance analysis (pertaining to different parameter experiments).  
**30 points**
3. Results visualized via silhouettes algorithm and interpreted.
  - a. Results from all experiments with varying clustering parameters represented with silhouettes.  
**20 points**
  - b. In addition to silhouettes algorithm, verification of the cluster correction against the class attribute in the data set.
    - i. Confusion table is required for each experiment performed.  
**20 points**
    - ii. Analysis of the clustering performance in terms of accuracy of grouping based on the confusion tables.  
**30 points**
4. Discussion on overall clustering performance, results and conclusion on the assigned clustering algorithm usability for the task.  
**20 points**
5. Indicate the accuracy, scalability and interpretability of the assigned algorithm.  
**10 points**

While concise reports are appreciated, the goal is to include all relevant information, particularly emphasizing your results discussion, interpretation of experiments as well as the final conclusion.

Single link hierarchical – Teams Ashley, Felicia, Joe, Melissa

Wards link hierarchical – Teams Josh, John, Theron, Zachary

Complete link hierarchical – Teams Robert, Wagner, Austin

EM clustering – Teams Jose, Ryan A., Allina, Tuvshintur