

Rubric: (accounts for 15% of the final grade.)

Choose data of your choice, (labelled data) 1%

Preprocessing and Data Cleansing 1%

Feature Engineering 2%

Use K-means, EM, Hierarchical clustering algorithms 3%

Perform Evaluations, 2%

Compare and decide which clustering result is the closest to the human labels 0.5%

Perform Error Analysis, 1.5%

Perform Visualizations, Graph the results 1%

One or Two pages of presentation 0.5%

Report, detail explanations 2.5%