Mini Project #2(B)

Task 4: MDS plots (numerical data dimensions only)

- (a) construct the data MDS plot (use the Euclidian distance) and visualize it via a scatterplot (use *metric* MDS – python sklearn.manifold.MDS)
- color the points by cluster ID (see task 3 in Lab 2(A))
- (b) construct the variables' MDS plot (use the (1-|correlation|) distance) and visualize it via a scatterplot (also here, use *metric* MDS)

Task 5: parallel coordinates plot (PCP)

- visualize the data in a parallel coordinates plot (all data dimensions, categorical and numerical)
- come up with a meaningful axes ordering by user interaction
- color the polylines by cluster ID (see task 3 in Lab 2(A))

Task 6: find a good PCP axes ordering from correlations

- numerical values only: use the correlations observed in the variables' MDS plot to help with the axis ordering – the user would click on points in sequence and the axes would be arranged in that sequence

Scoring and Due Dates

Each (task) bullet item carries 10 points

- an extra 10 pts for overall elegant implementation and function

Don't forget to

- label the axes and tick marks where appropriate
- show color legends where appropriate
- provide a meaningful header on each plot

Due date

- due April 1, end of day
- no group project, single only

Deliverables

Submit on Brightspace

- voice-narrated video file to show all features of your software in action
- in the video discuss any interesting observations you were able to make in the data
- 2-3 page report
  - describe interesting observations (beyond the video)
  - mention anything noteworthy about implementation (beyond the video)
- zip file with complete source code as well as the data
- submit video as an extra file

Grading

Grading

- TA will pick students at random for thorough code review sessions
- you better know your code !!!
- so, please do not just copy code beyond the D3 templates
- or even worse, videotape someone else's program