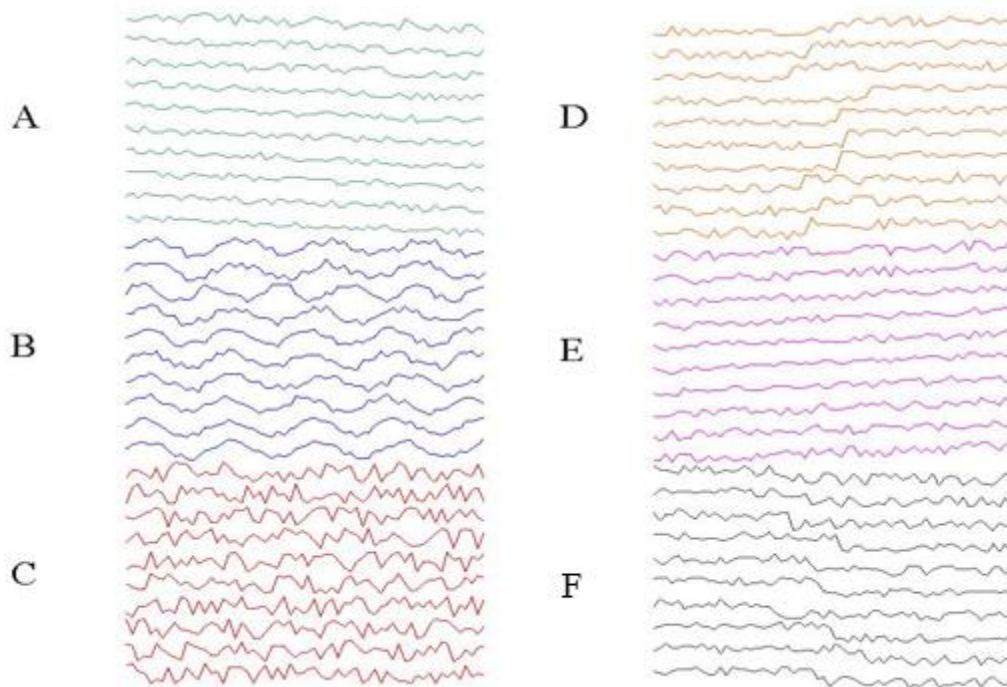


CSCD 429/529 Data Mining Lab 6 (20 points)

Clustering the control charts

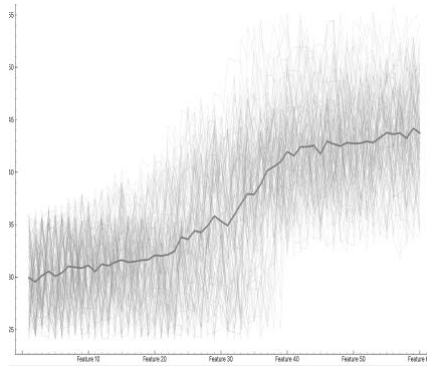
- **Data Description:** The dataset *synthetic_control_data.csv* contains 600 examples of control chart time series data. The data is stored in an ASCII file, 600 rows, 60 columns, with a single chart per line. There are six different classes of control charts:
 - Normal
 - Cyclic
 - Increasing trend
 - Decreasing trend
 - Upward shift
 - Downward shift

The following image shows ten examples from each class: (A) Downward Trend. (B) Cyclic. (C) Normal. (D) Upward Shift. (E) Upward Trend. (F) Downward Shift.

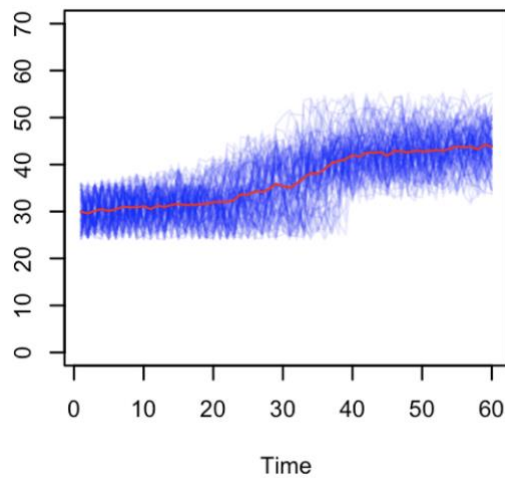


Task Description:

- 1) **Clustering and Visualization using RapidMiner:** You can use any clustering methods available in Orange to generate the six clusters from control chart data, then use the visualization tool provided by Orange to show the clusters. As an example, the following graph is one of the clusters generated by Orange.



- 2) **Clustering and Visualization using R:** You can use any clustering methods available in R to generate the six clusters from control chart data, then use R to visualize the clusters. As an example, the following graph is one of the clusters generated by R.



Deliverables:

- (10 points) The diagram of Orange process and six cluster images for Task 1.
- (10 points) Workable R code and six cluster images for Task 2.
- Include everything in a pdf file and submit your file via Canvas.