

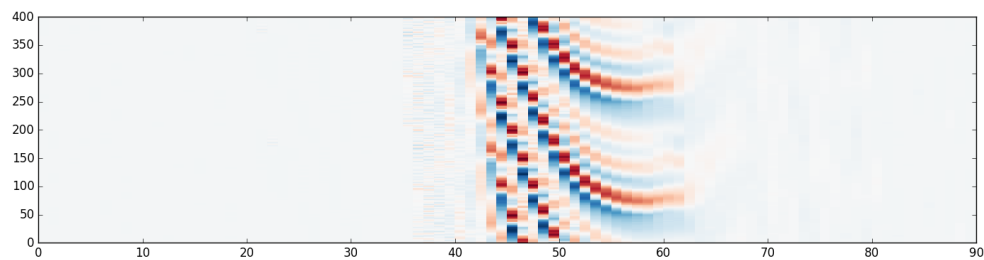
Speech and Speaker Recognition DT2119

"Lab1: Feature extraction"

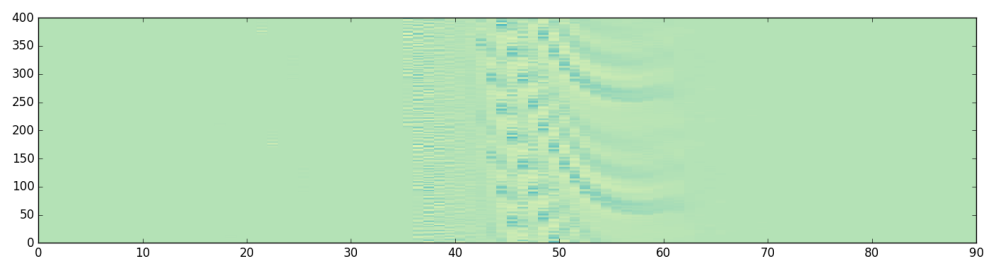
April 7, 2017

4 Mel Frequency Cepstrum Coefficients

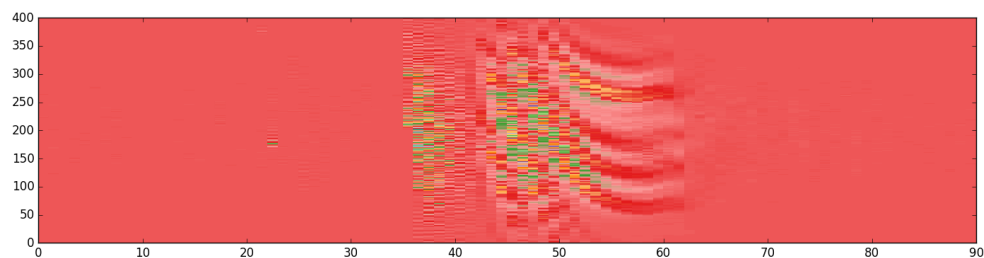
(a) Enframe



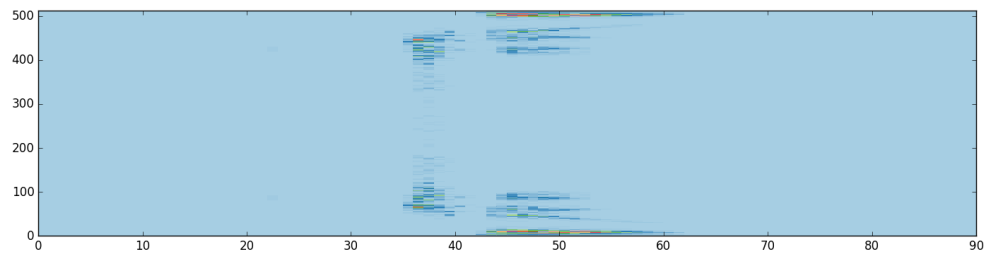
(b) Pre-emphasis



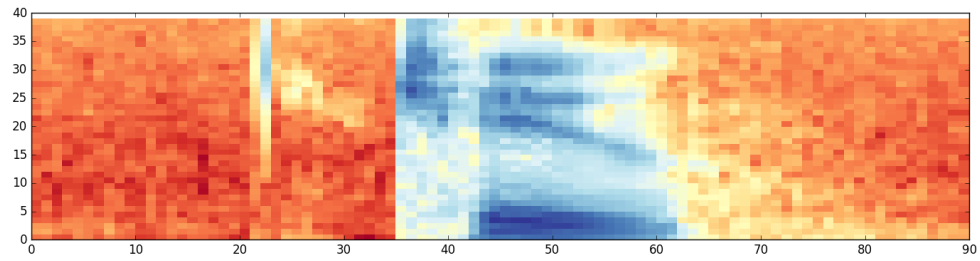
(c) Hamming Window



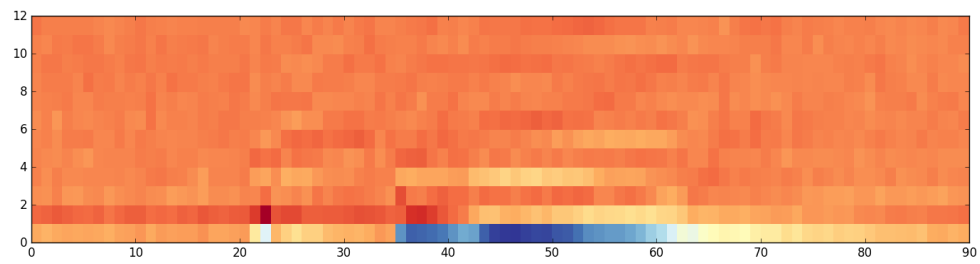
(a) Fast Fourier Transform



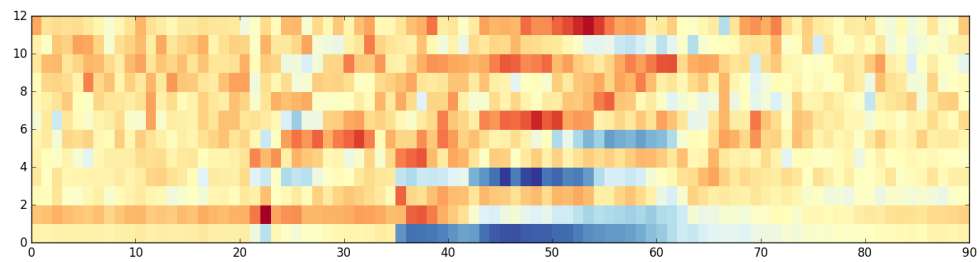
(b) Mel filterbank log spectrum



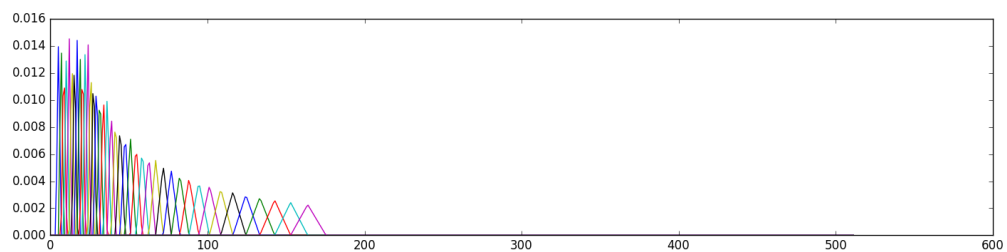
(c) Cosine Transform



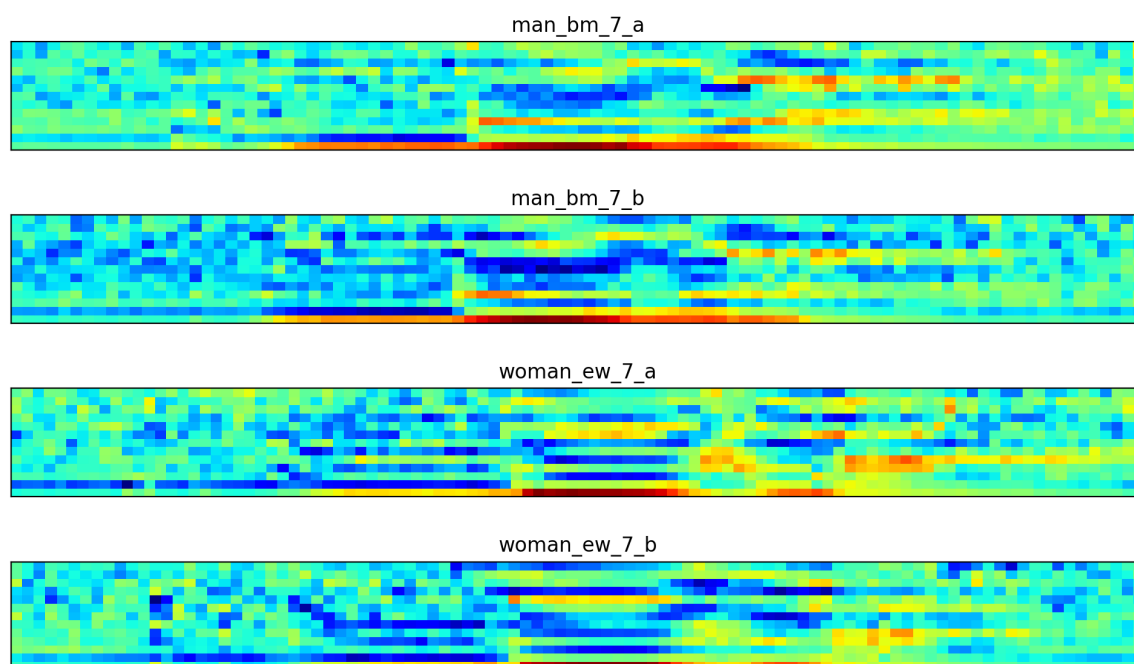
(d) Cosine Transform and Liftering



(e) Bank of triangular filters in the Mel frequency scale



(a) MFCC for different utterances



5 Feature Correlation

Figure 4: MFCC correlation

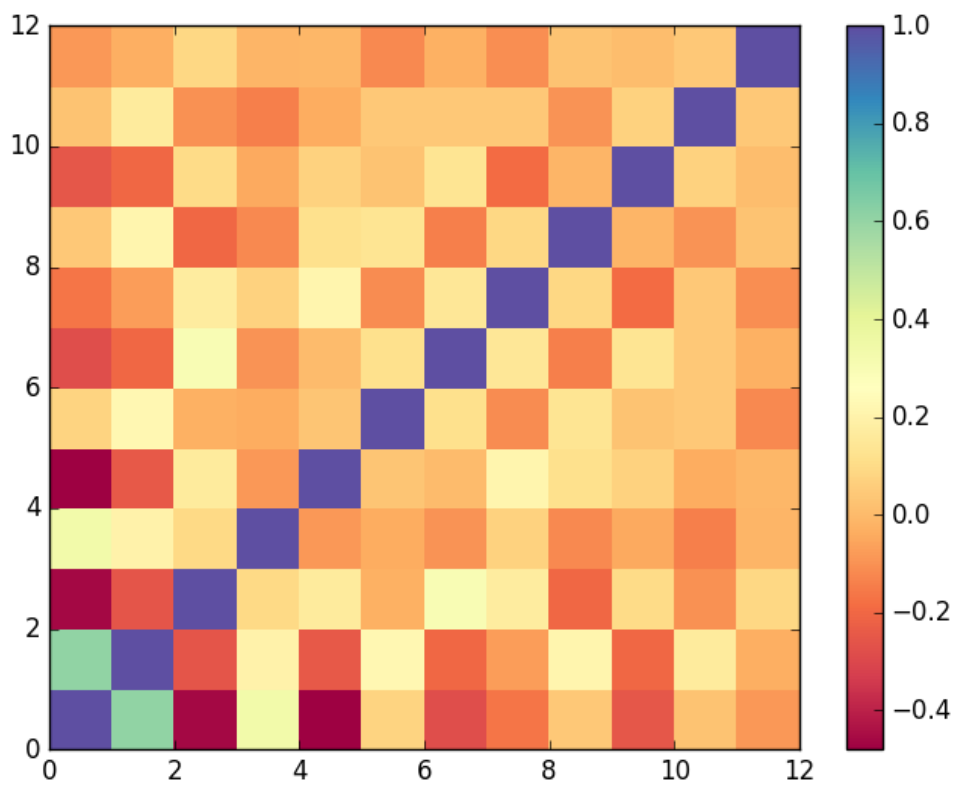
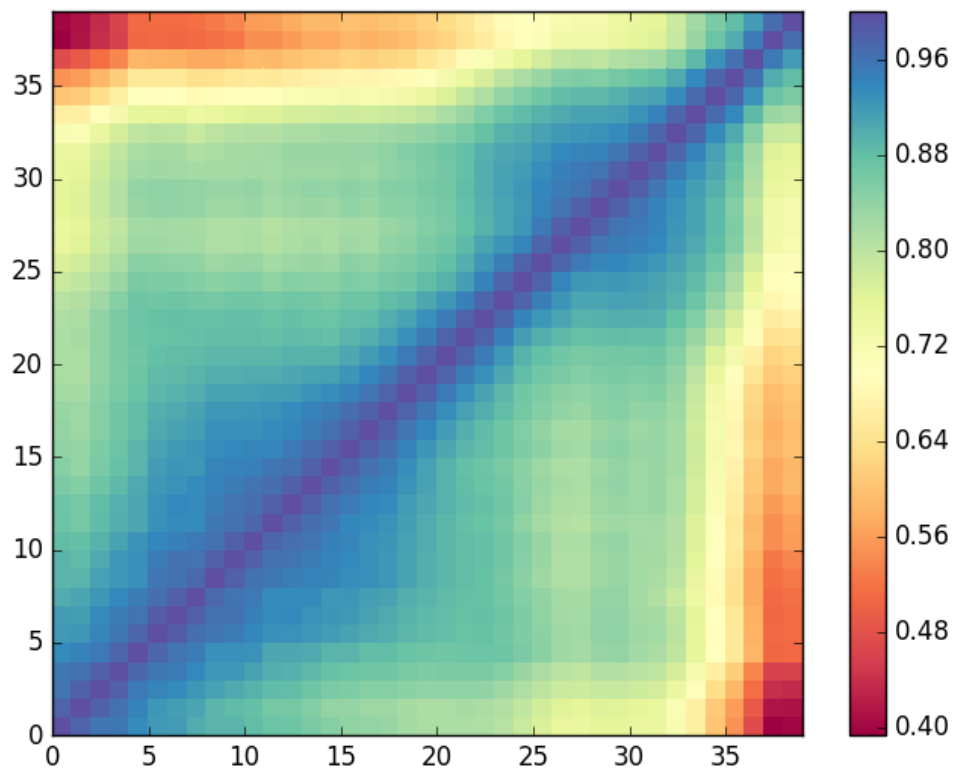
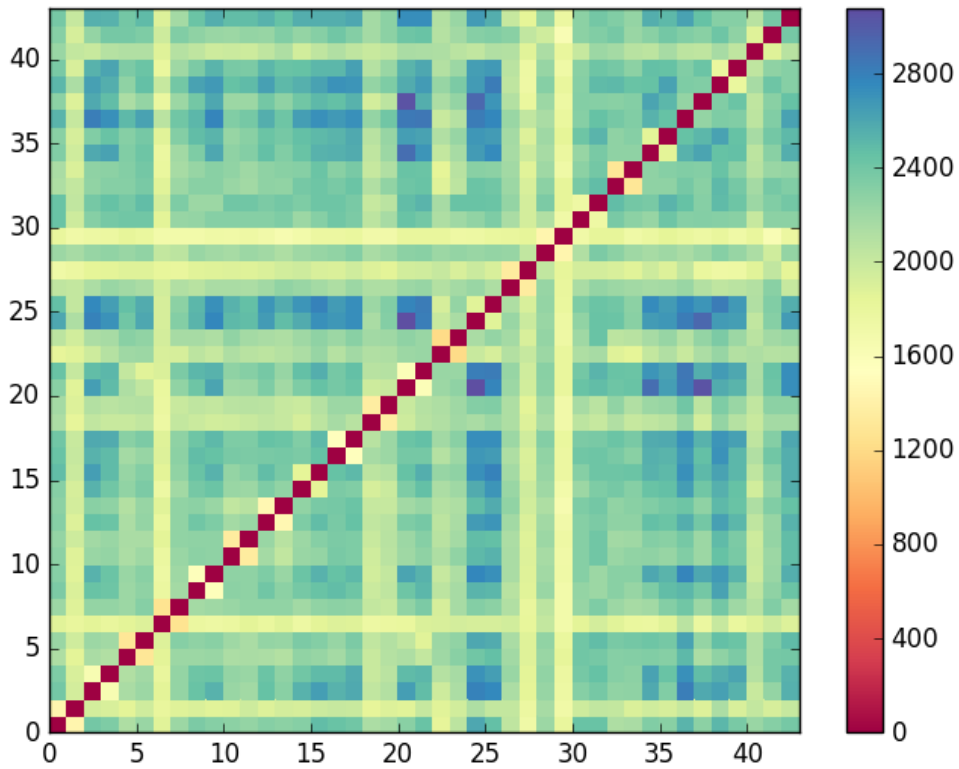


Figure 5: Mel filterbank correlation

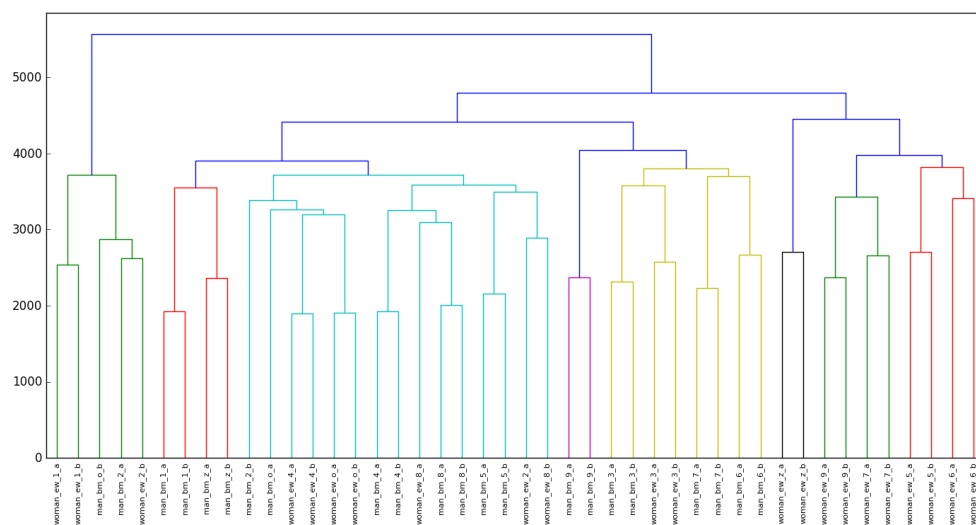


6 Comparing Utterances

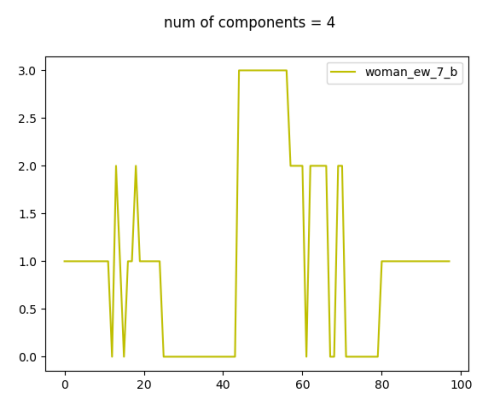
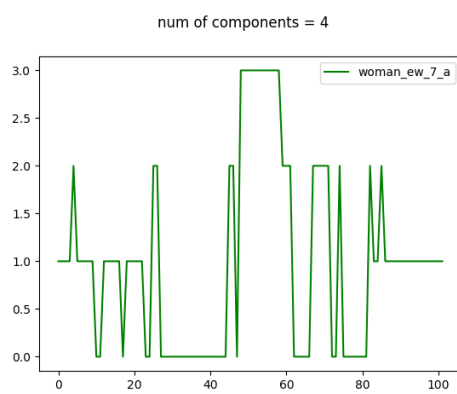
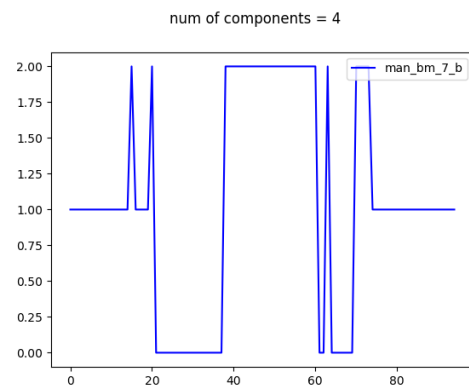
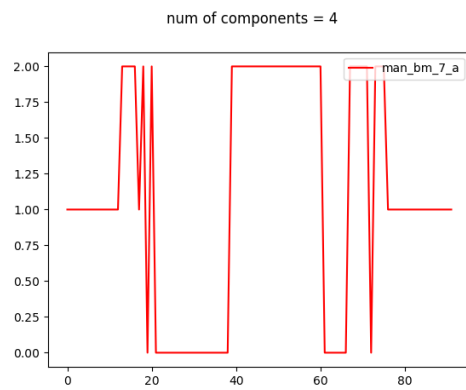
(a) Global Distances between utterances



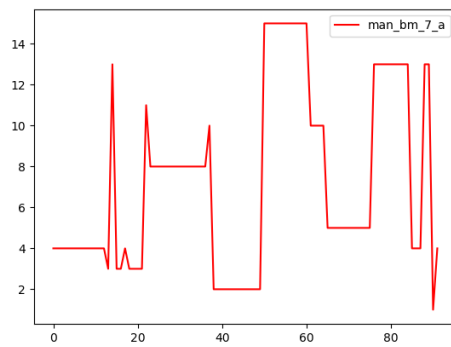
(b) Dendrogram



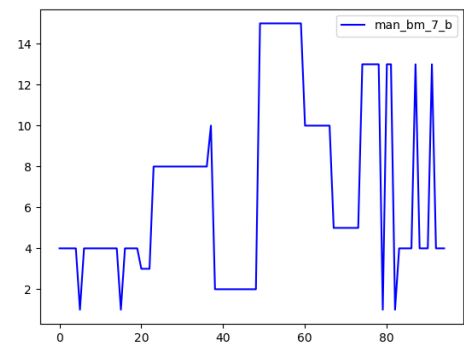
7 Explore Speech Segments with Clustering



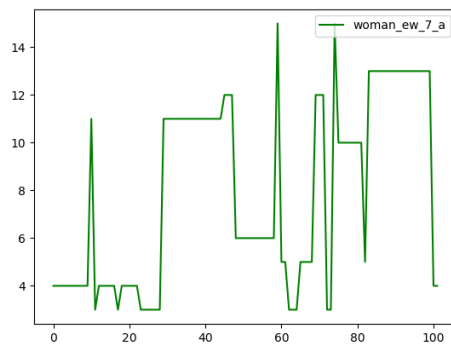
num of components = 16



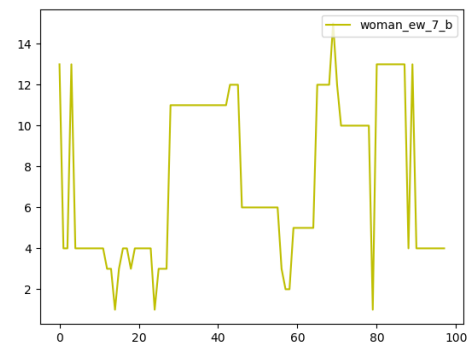
num of components = 16



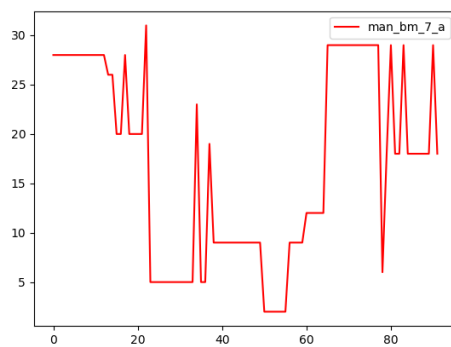
num of components = 16



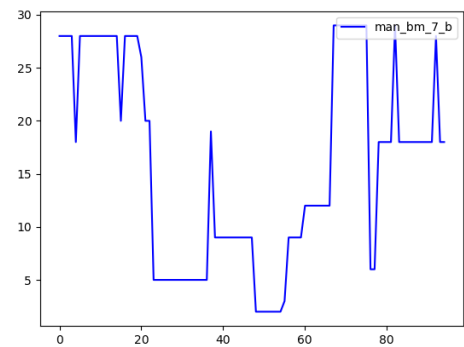
num of components = 16



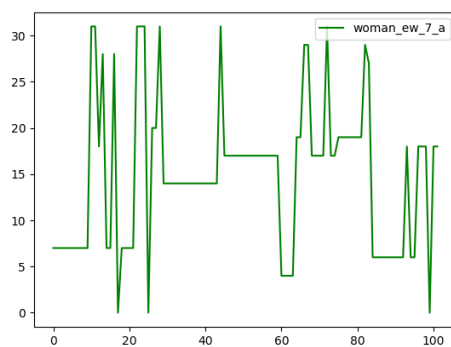
num of components = 32



num of components = 32



num of components = 32



num of components = 32

