## Steps and explication

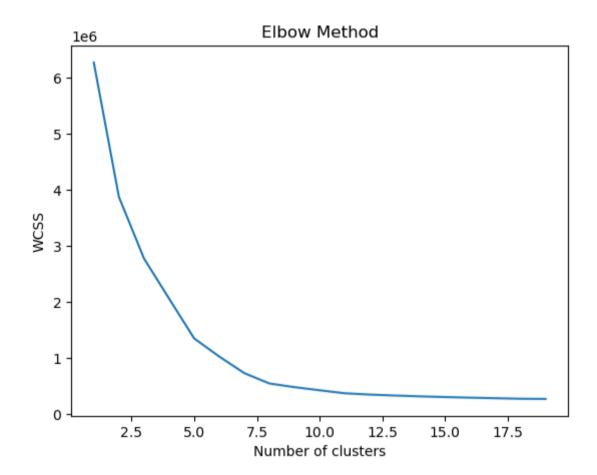
Step 1: Transpose the data into shape (999,18)

```
2
                                                             9
    62.44 59.26 59.26
                       59.26 59.26 59.26
                                          0.00
                                                 0.00
                                                       0.00
                                                            0.0
    71.62
          73.12 67.04
                       60.75 60.75 60.75
                                          0.00
                                                 0.00
                                                       0.00
                                                            0.0
                                         0.00
                      70.39 70.39
    71.10
          70.39
                70.39
                                   70.39
                                                 0.00
                                                       0.00
                                                            0.0
    64.74 73.51 73.51
                       73.51 73.51
                                   73.51
                                          0.00
                                                 0.00
                                                       0.00
                                                            0.0
                      84.72 84.72 84.72
    82.62 72.98 79.20
                                         0.00
                                                0.00
                                                       0.00 0.0
     ...
           ...
                 ...
                        . . .
                               . . .
                                    . . .
                                           . . .
                                                 ...
                                                       . . .
                                                            . . .
994 96.24 79.78 86.10 103.12 93.60 93.60 93.60 93.60
                                                     93.60
                                                            0.0
995 91.84 84.39 78.61 103.02 84.56 78.39 78.39
                                               78.39
                                                      78.39
                                                            0.0
996 94.33 86.56 90.42
                      96.51 96.51
                                   96.51 96.51
                                               96.51
                                                     96.51
                                                            0.0
                      83.58 74.42 82.61 82.61
997
    85.70 91.19
                77.48
                                               82.61
                                                     82.61
                      99.95 99.95 99.95 99.95 99.95 0.0
998 93.74 90.29 99.95
     11
         12
             13
                 14
                      15
                          16
                               17
    0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 0.0 0.0
        . . .
             . . .
                 . . .
                     . . .
994 0.0
        0.0
             0.0
                0.0 0.0
                          0.0
995
   0.0 0.0 0.0 0.0 0.0 0.0
996 0.0 0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 0.0 0.0
998 0.0 0.0 0.0 0.0 0.0 0.0 0.0
[999 rows x 18 columns]
```

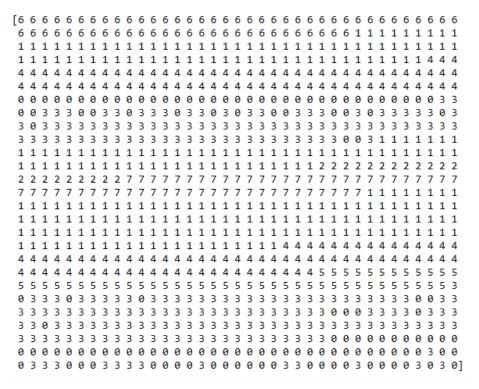
\_\_\_\_\_

Step 2: Apply PCA

Step 3: Using Elbow method, I choose number of clusters = 8



**Step 3**: Feed the data to the KMeans(n = 8). We got a list of cluster label for each column in the merged\_file like this: shape(999,)



**Step 4**: We will go through each file, plot the average value of each column in each file. The color will correspond to the label in the list (I update the position of label after go through each file). I am a bit

confused about the meaning of color , but I thought that the KMeans only able to say whether this group of data belong to which session because our data have nothing relate to emotion . That is the reason I put the legend as session, for example here the KMeans indicate that the blue data belong to session 1 and violet belong to session 4 .

