Partitioning Two-Thousand Stocks by their return series

Attached .csv file contains return series of 2,000 stocks(columns) for 3,000 days(rows). Please make 10 partition groups of 2,000 stocks satisfying following conditions.

1. Each partition group has 200 stocks.
2. Minimize inter-group return correlation and maximize intra-group return correlation.

   Objective function will be CARS/TICS.
3. Intra-group return correlation TICS is defined as below.

   For each partition group with label k, 200x200 correlation matrix IC(k) can be calculated. ICS(k) is a summation of all elements in IC(k).

   TICS = ICS(1) + ICS(2) + … + ICS(10)
4. Inter-group return correlation CARS is defines as below.

   AR(k) is a return series obtained by averaging over 200 stocks' return series. Then AR(k) is a vector with 3,000 elements(3,000 days). CAR is 10x10 correlation matrix of [AR(1), AR(2), … AR(10)] and CARS is a summation of all elements in CAR matrix.
5. Correlation matrix is defined in:

   https://en.wikipedia.org/wiki/Covariance_matrix#Correlation_matrix

Please submit answer in the format of .csv which maps each stock(column) to one of 10 groups as below: {stock_index, gropu_id}. Also please kindly submit your source code or pseudo-code and descriptions.

0,0
1,0
2,1
3,1
…
1998,5
1999,1
<end of file>

➔ 0^{th}, 1^{st}, … stocks are in group-0; 2^{nd}, 3^{rd}, …, 1999^{th} stocks are in group-1.

- Example

"hfc_20170614_example.csv" file has 4 stocks(columns) and 40 days(rows). If we partition four stocks into two groups to satisfy above conditions, CARS/TICS is minimized when 0^{th} and 2^{nd} stocks are in one partition(group_id = 0) while 1^{st} and 3^{rd} stocks are in another partition(group_id = 1) as below. And CARS = 1.740546 and TICS = 4.15635

- Answer for the example

0,0
1,1
2,0
3,1
<end of file>

END