**Training Data Set Creation**

The training data set used for this analysis is shown below (first 5 lines):

Dimensions of the final data set:

# of clusters created: