Assumptions:

1. Clustering is based on the distance between two points. However, these points should be in the same timestamp. We divided the data into many parts using the timestamp, and later clustered the data within all the values of that timestamp. The cluster index is initially set to 1, and then gradually incremented.

Description:

1. calculate_distance is used for calculation of distance using 4 variables where x1,y1 and x2, y2 x axis and y axis position of two points respectively.

2. This keyReturner accepts the following inputs: a dictionary containing data points and their corresponding positions, and the current x and y coordinates. It then determines the key within the dictionary where the distance between the position provided and the stored position is the minimum. This ensures that the related data points are assigned to the same cluster. The function also returns the key of the cluster to enable subsequent data to be assigned to the same cluster.

3. func1 uses current_timestamp and data_dictionary and uses function keyReturner to create a cluster then append the data to the newly created csv using writeToFile and append_data_to_csv functions.

4. The last part of the code is used for reading the given CSV. then send it to multiple functions created for the compiling the new CSV i.e data.csv.