Assignment 8

You have a Eco-hotel Data Set (https://archive.ics.uci.edu/ml/datasets/Eco-hotel) in the format of comma separated file (dataset-CalheirosMoroRita-2017.csv), it can be read and open in excel easily.

It is required that you observe the customer reviews sentiment of each syntax by weighing certain words. For example, the third row has the following syntax

"The hotel it is fantastic built by the sea, living together with nature. Environment it is great as well as people and service.

We full enjoyed the place, and facilities.

Thanks for the ""cidreira"" and ""madalenas"" tea at reception"

Here the following keywords are important

fantastic (the hotel it is fantastic ....) 2

nature (living together with nature) 1

great (Environment is great) 3

enjoy (we full enjoyed the place) 1

thank 1

----------------------------------------------

Sentimental Score : (8/39)\*100 = ceil(20.51) = 21

Where 8 is the total weight of these keywords

39 is the total words in this syntax

Another example is

"One dream! Cozy and comfortable Hotel! The best personalized one! All the employees are congratulated! Very nice taste"

Cozy 1

comfortable (Cozy and comfortable Hotel) 1

best (the best personalized one!) 2

very nice (Very nice taste) 3

----------------------------------------------------------------

Sentimental Score : (7/18)\*100 = ceil(38.88) = 39

Your job is to define a map for these words first, for example if you have only two syntax as shown above the map will be shown as follows

fantastic:2

nature:1

great:3

enjoy:1

thank:1

Cozy:1

comfortable:1

best:2

very nice:3

You will use this map to calculate the sentimental score, as shown in the above two example. Now, your job is to find these words (mostly these words are common in all the syntax, so if you go through only 100 syntax that would be enough).

After finding the keywords and make a map out of it. Then you can calculate the sentimental score of all the syntax, with the help of this map. You will store it into two-dimensional list such that the first column is syntax reference and second is its corresponding score (you can use any other container like dictionary etc.)

for example

List:

Syntax-1 : 21

Syntax-2 : 39

etc.