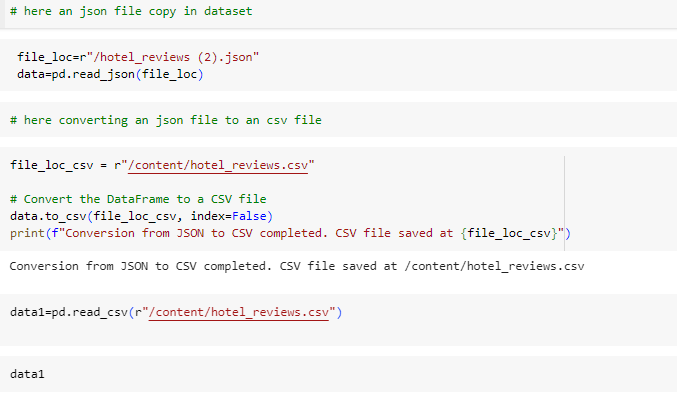
PROBLEM STATEMENT:Each hotel review is an the text format as json file.

Convert text to an vector form and to extract an feature from the data

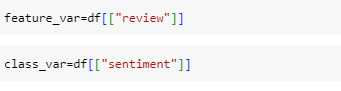
Convert:json file to an csvfile



* 2) Data Exploration:
* In the data am checked an the it has an two colums
* 1.review 2. **Sentiment**



And we are based on the problem statement we are devided in a



In an eda I checked data, there is an no html tags,in the data having no lower case ,no url , it has lot of emojis and having a lots of special charactors are there



3) Text Preprocessing:

1.converting an text in a lower case

2.remove an all special charactors inan reviews

3.converting an emoji to an demoji as text

4.remove an stop wors

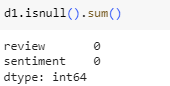
5.lemmatization technique in an English word   


6.and join the all reviews in on single 



And again check any null values and any duplicate values





  4)  Data Exploration:

