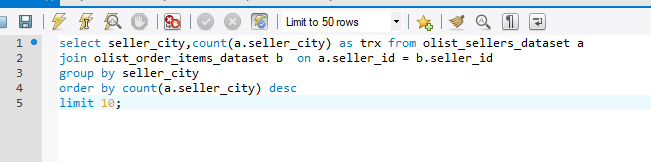
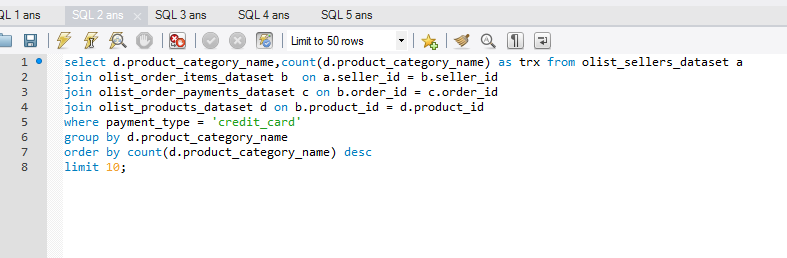
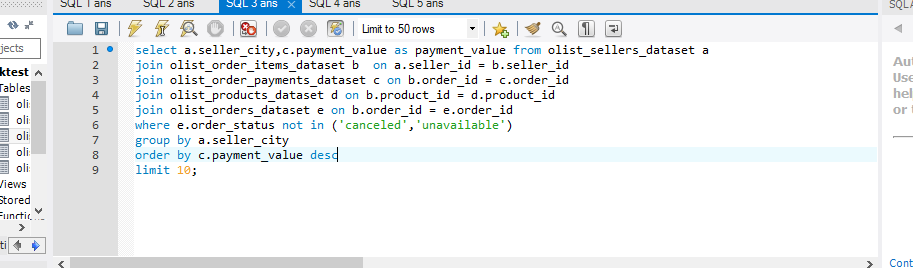
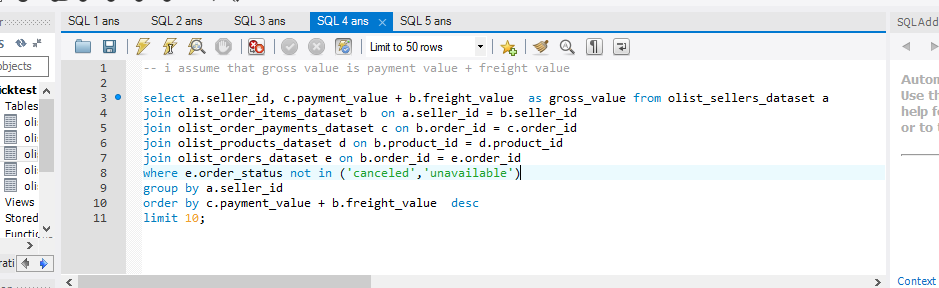
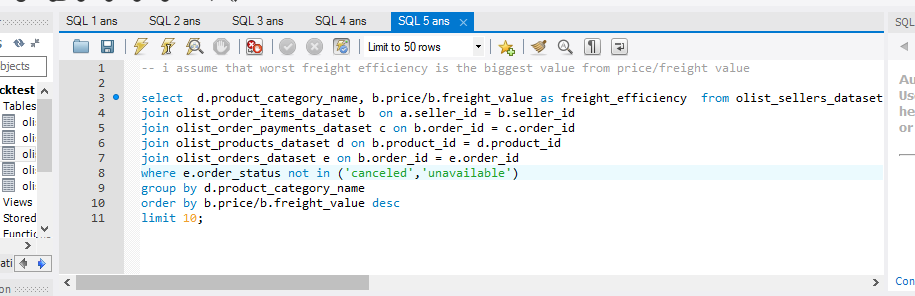
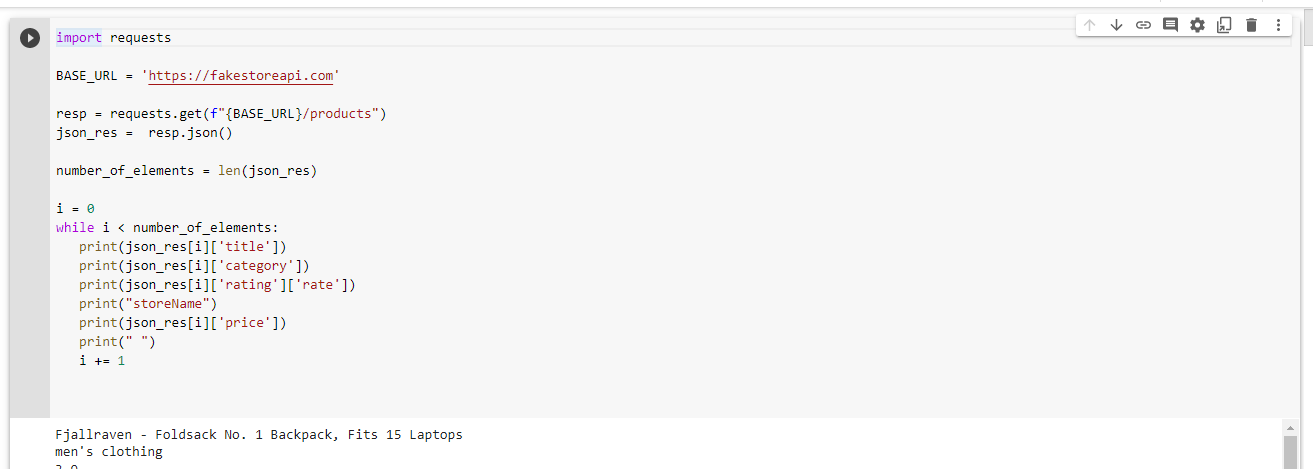
**SQL**

1. 
2. 
3. 
4. 
5. 

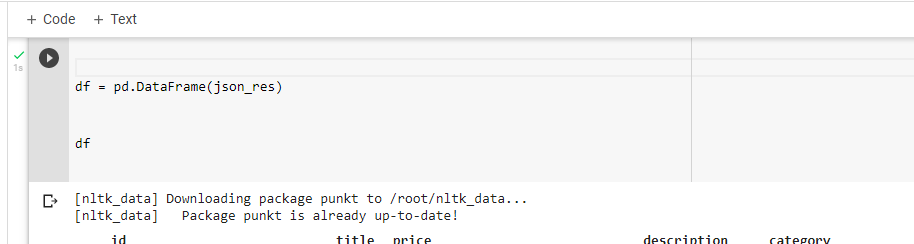
**Data Processing**

i'm so sorry that i can't to find a proper API dataset. so i use FakerAPI store to be process

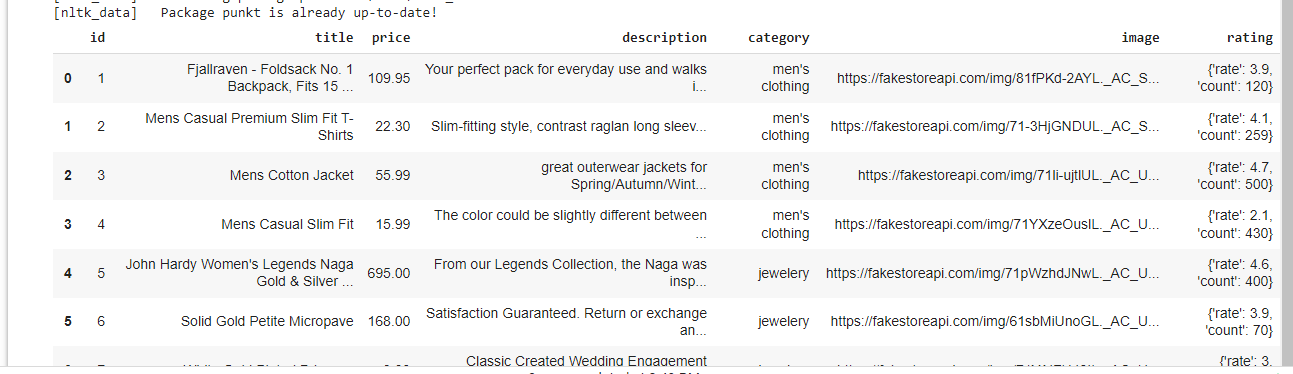
this how I processing the data

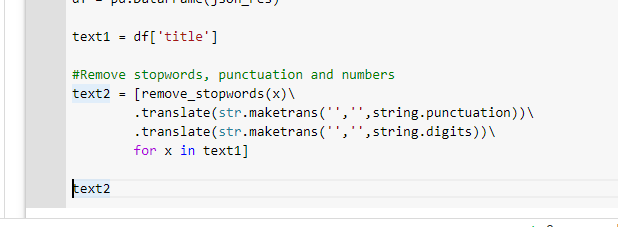


**Machine Learning**

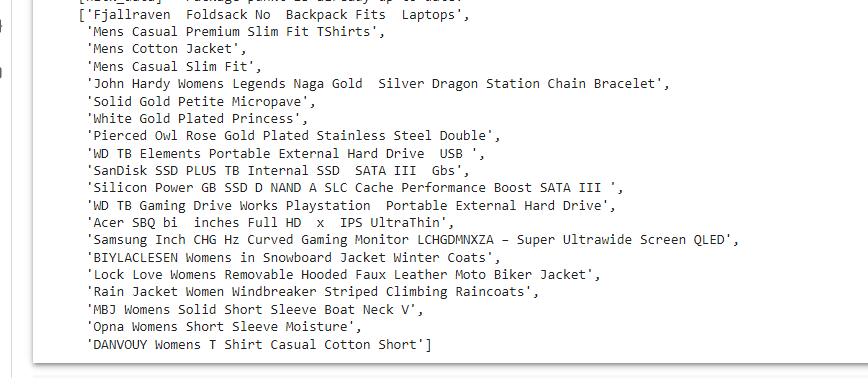
 

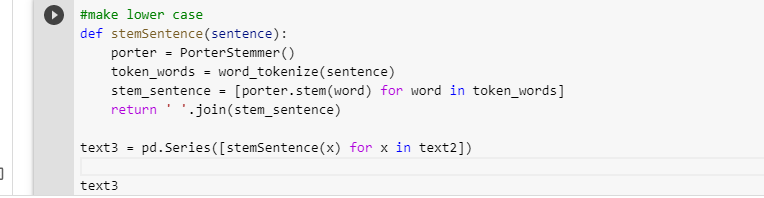
in this part I generate the data from API end then I put on dataFrame

the result: 

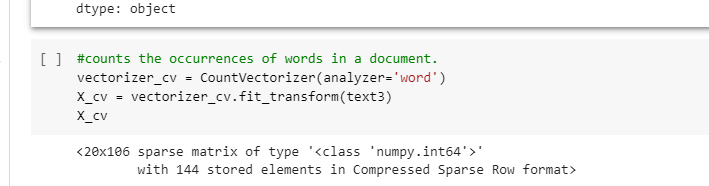


In this part I take 1 part of column to become benchmark and then I cleansing the data with stopwords, punctuation and numbers

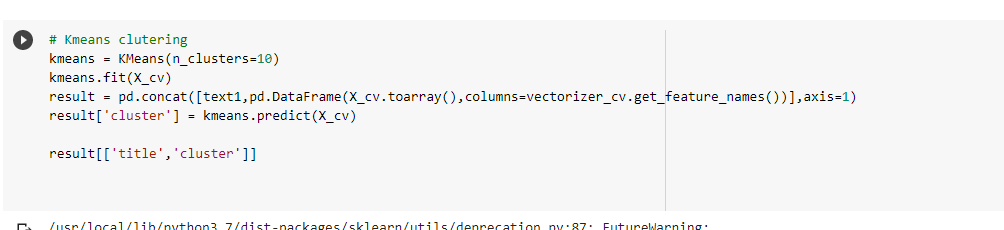
The result: 



In this part I make all the words become lower case, to purpose become uniform



In this part I count the occurences words in the data.



The last part is I use Kmeans clustering to clustering the data

So, the product name is have a cluster

Result: 