**Minor Project : Text Summarization**

**Introduction**

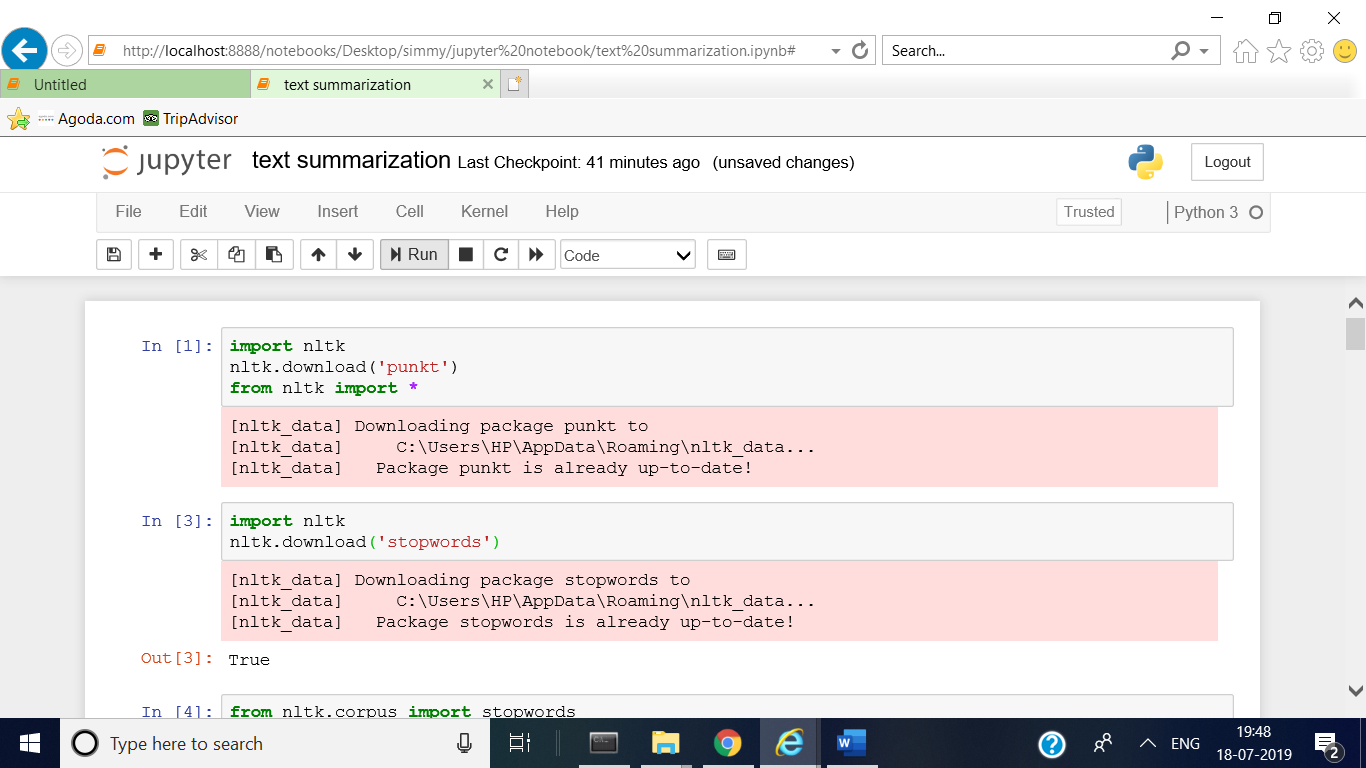
Text summarization is a subdomain of [Natural Language Processing](https://stackabuse.com/what-is-natural-language-processing/) (NLP) that deals with extracting summaries from huge chunks of texts. *Text summarization is the process of distilling the most important information from a source to produce an abridged version for a particular user and task .*

**Objective**

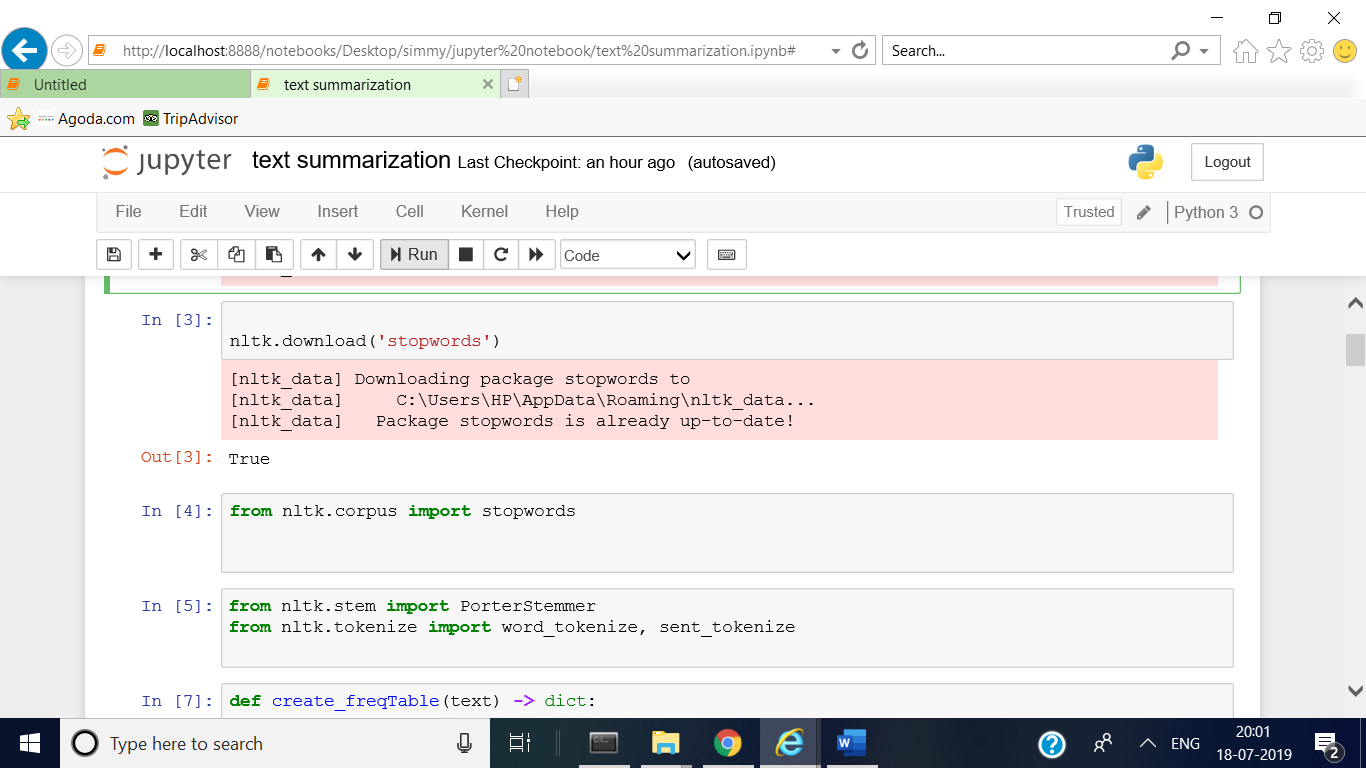
The main objective of a text summarization system is to identify the most important information from the given text and present it to the end users.

**STEP 1** Import important libraries.

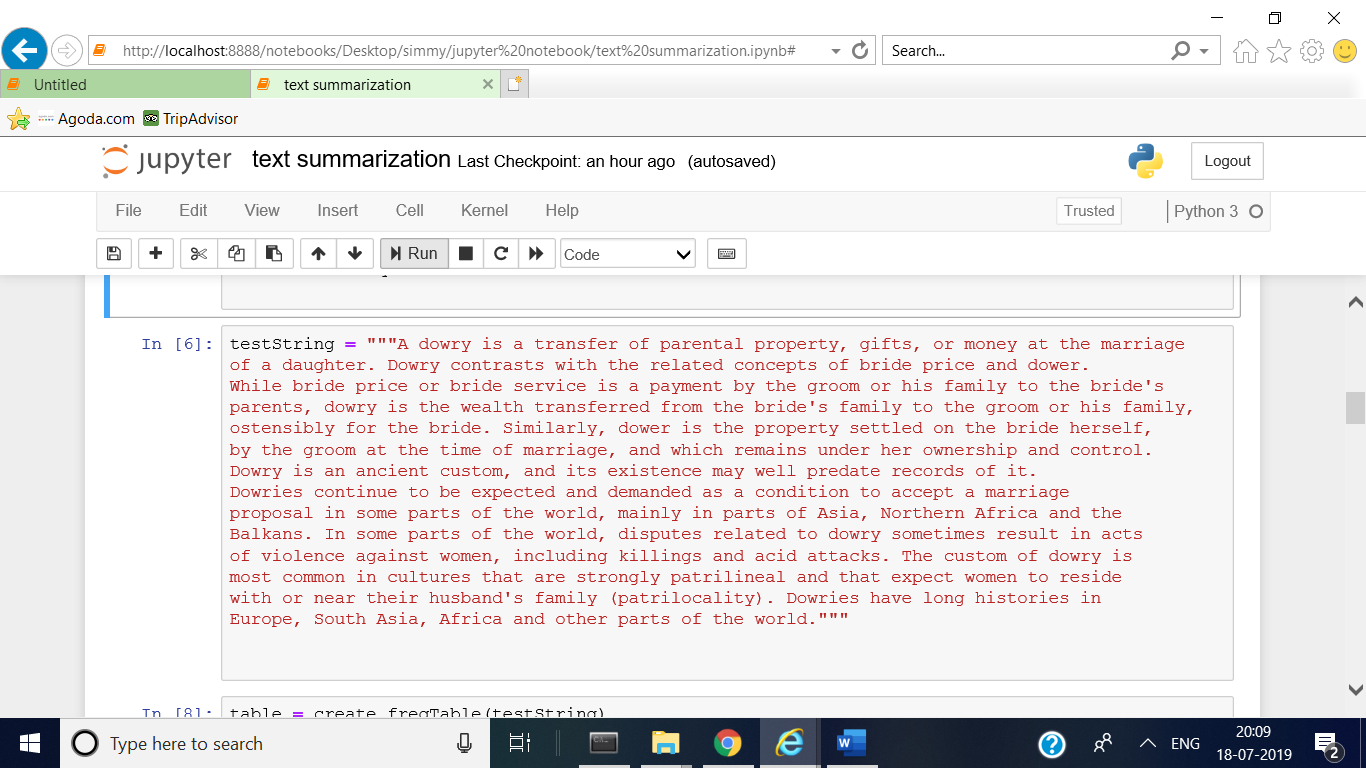
Firstly, import all the libraries that are required. Most important is nltk



Nltk has many dataset available to use. I downloaded stopwords, PorterStemmer, word\_tokenize and sent\_tokenize.

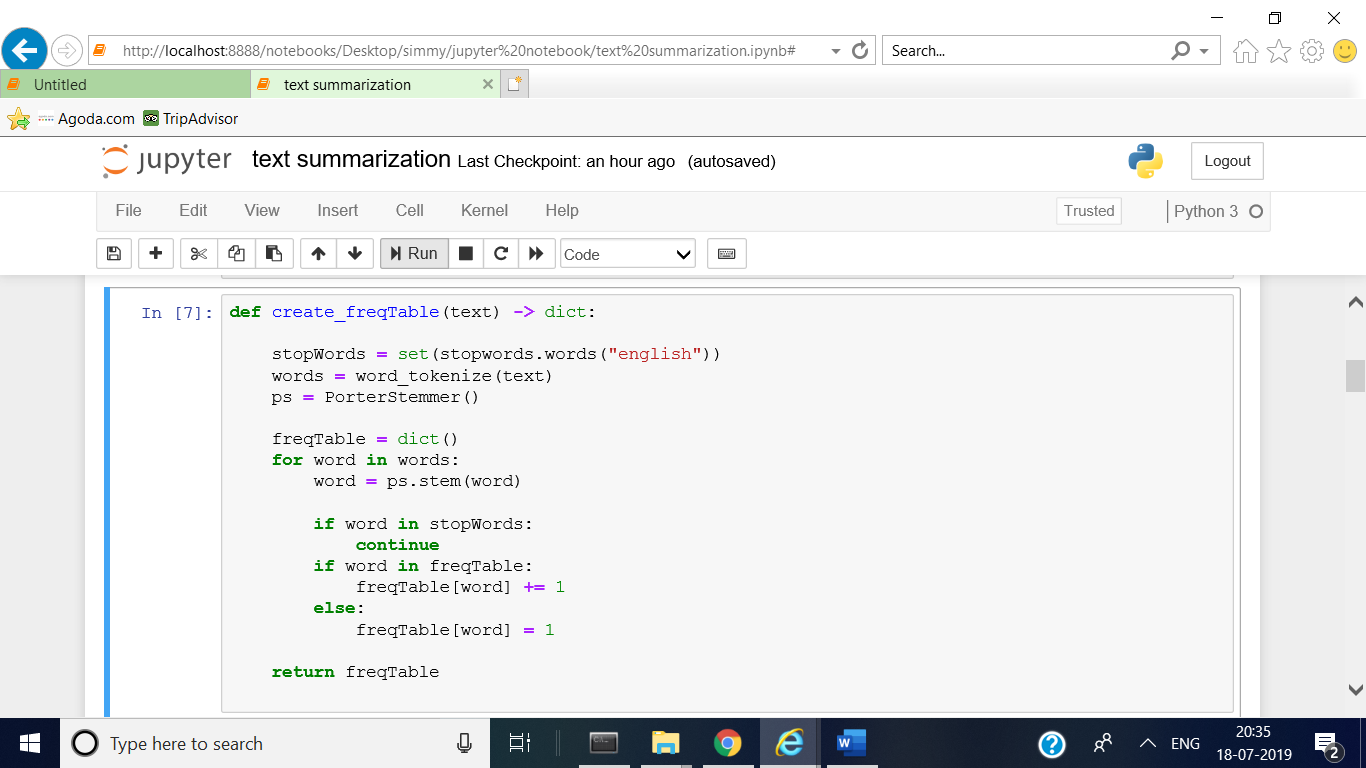


This is a test string . I’ll use this string from this step onwords.

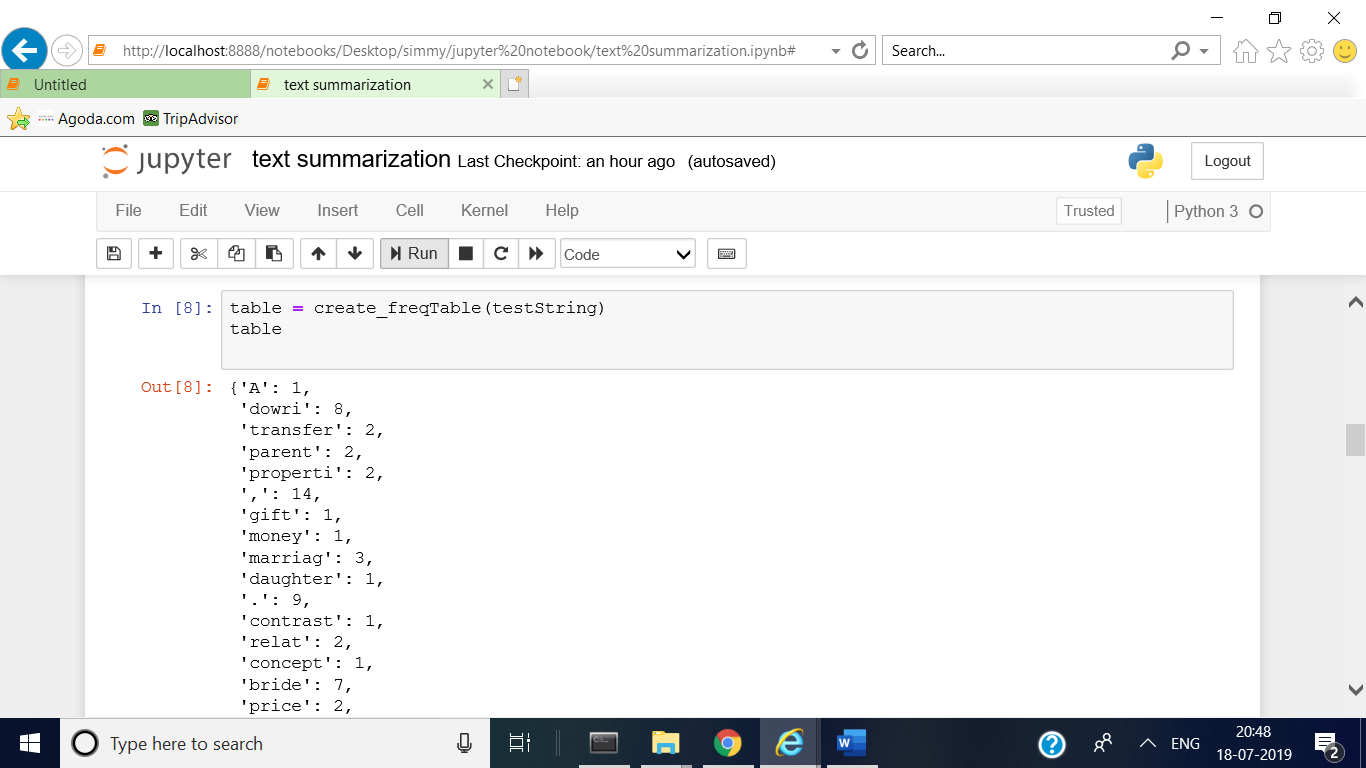


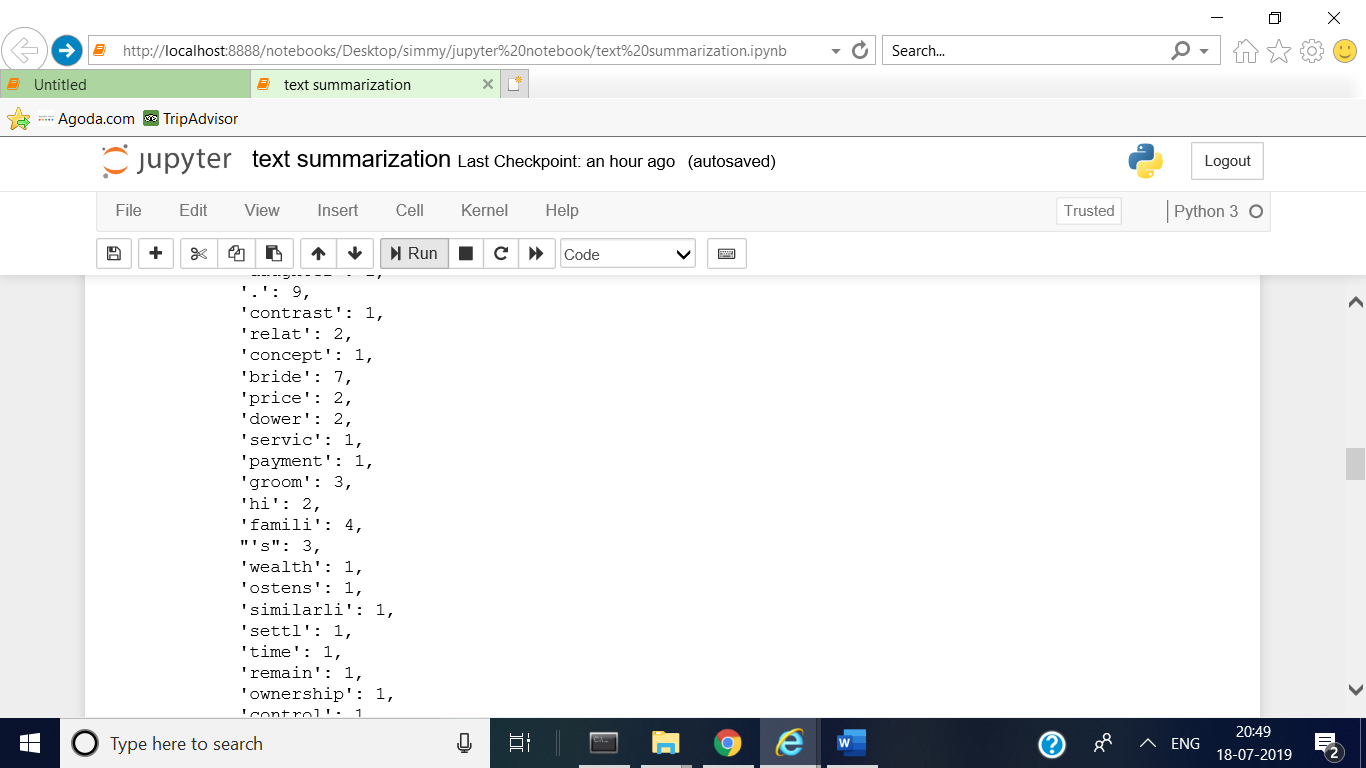
**STEP 2 :** create frequency table

A dictionary is created to calculate frequency of each word

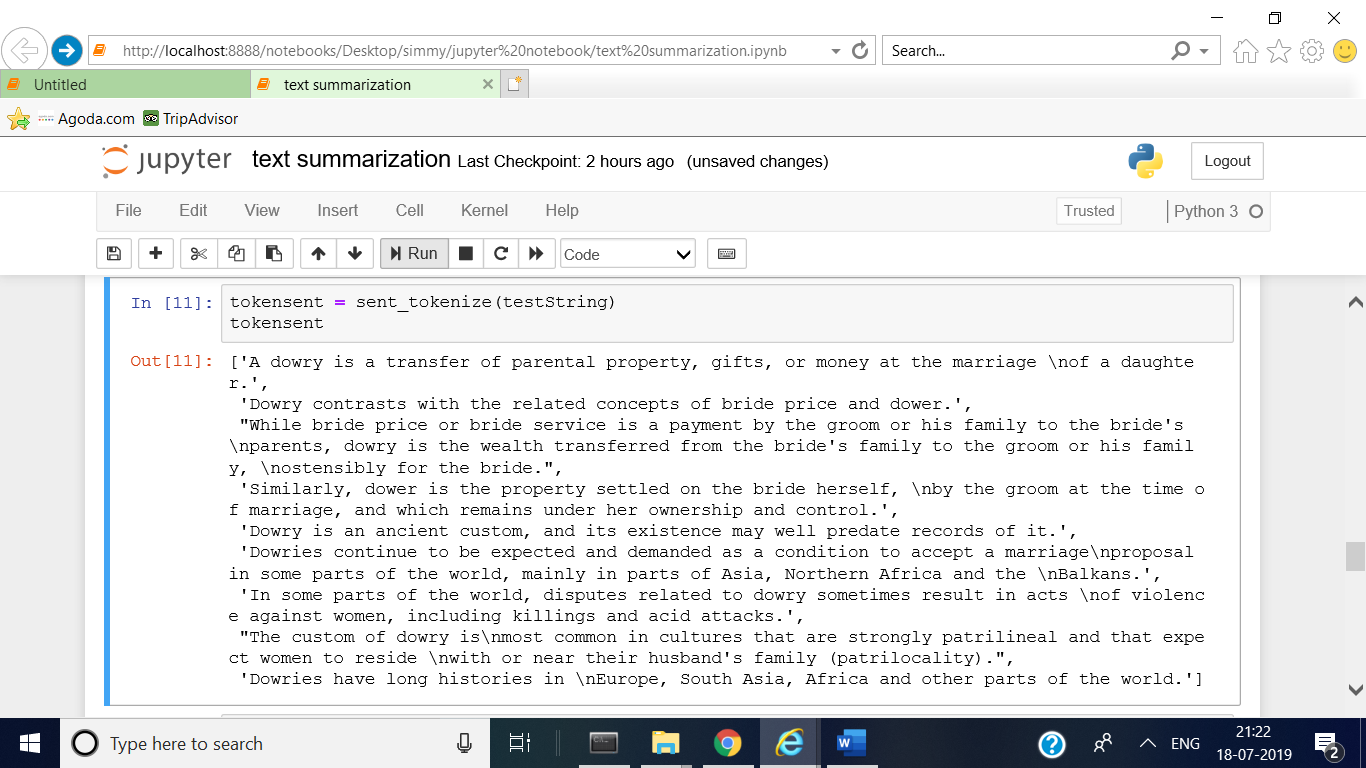


The output will look like.

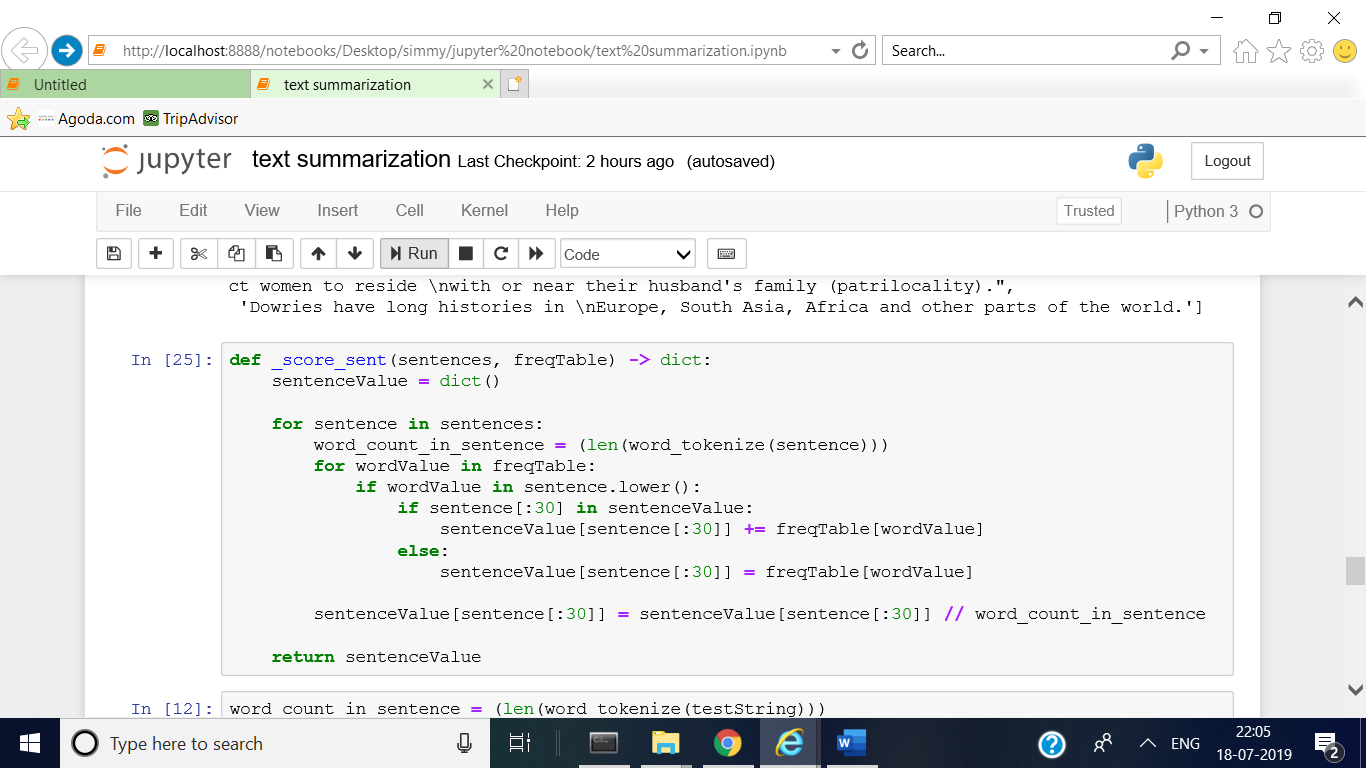




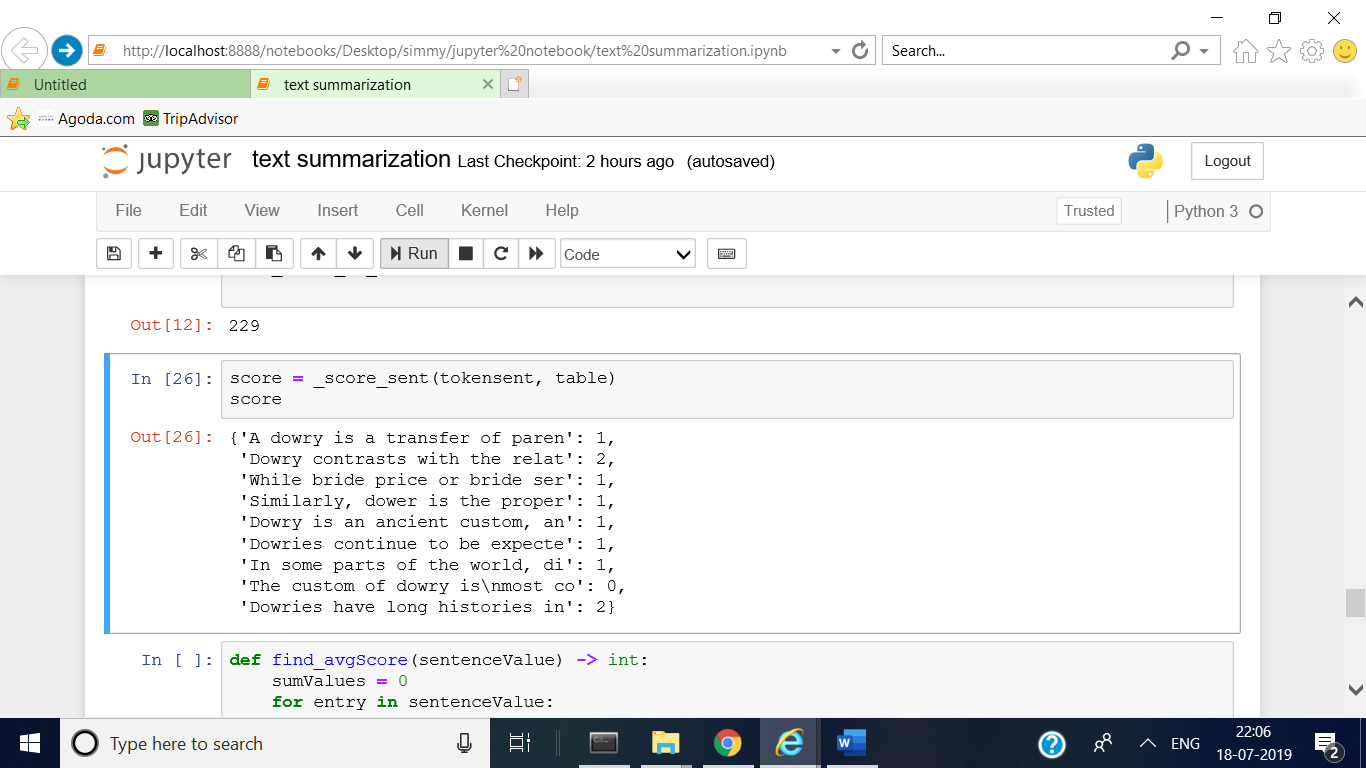
**Step 3:** Sentence tokenization :- tokenize the test string into sentences.



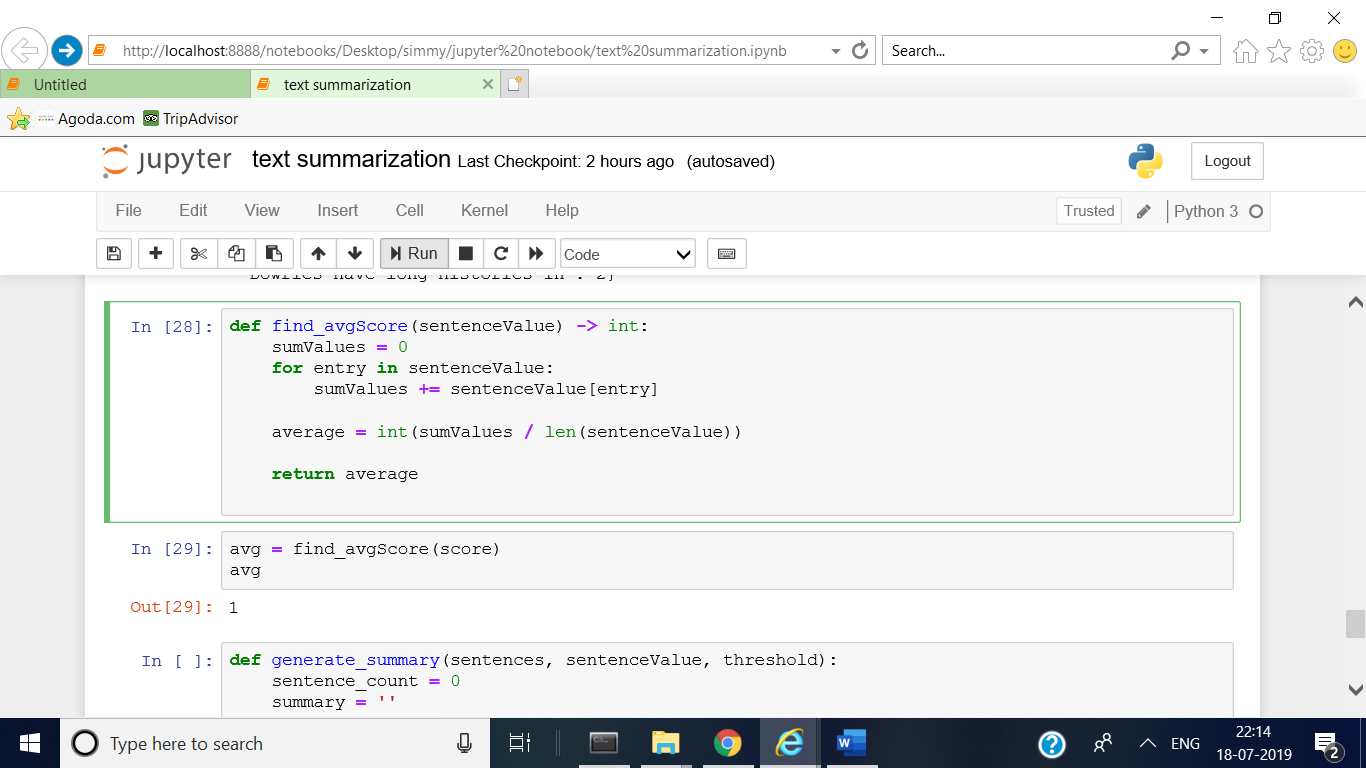
**Step 4:** score the sentences



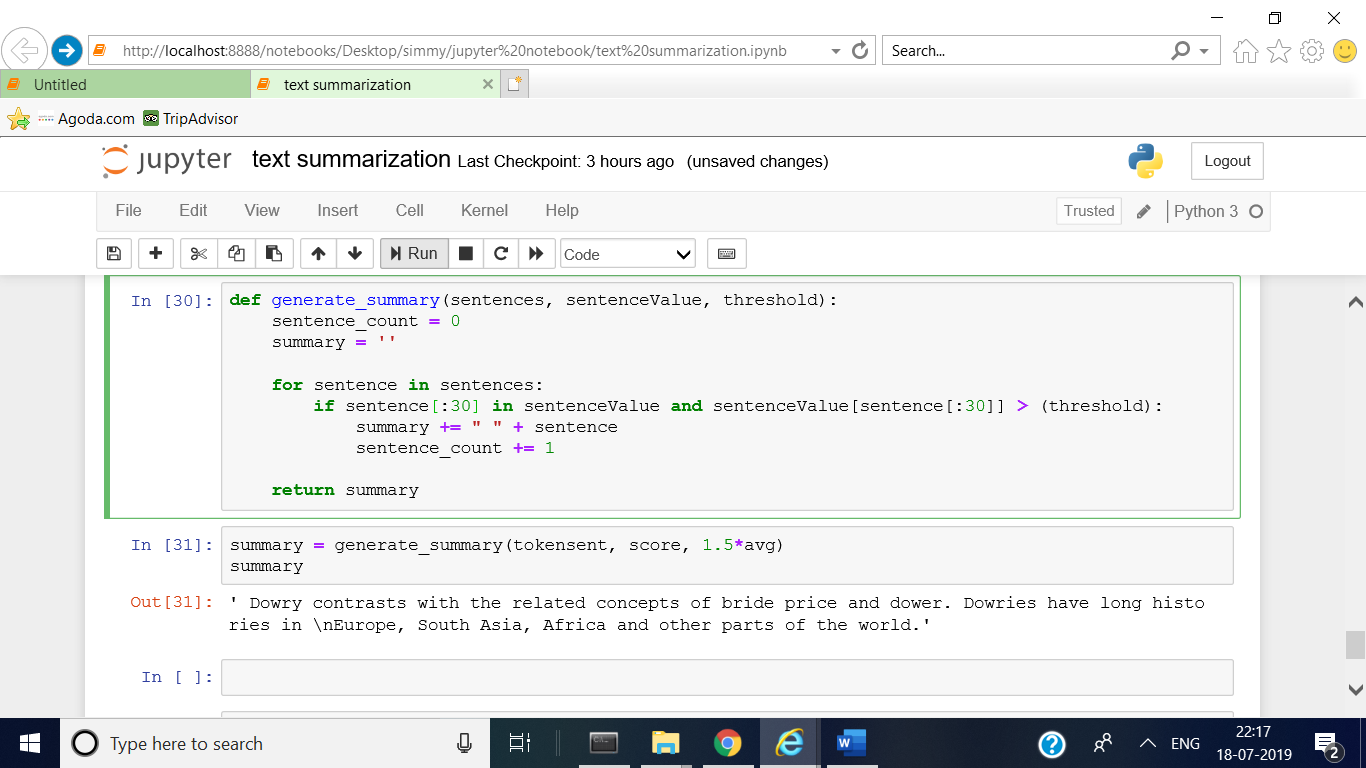
Output will look like:



**Step 5:** a function that will find average score (threshold) by dividing it by length of sentence value

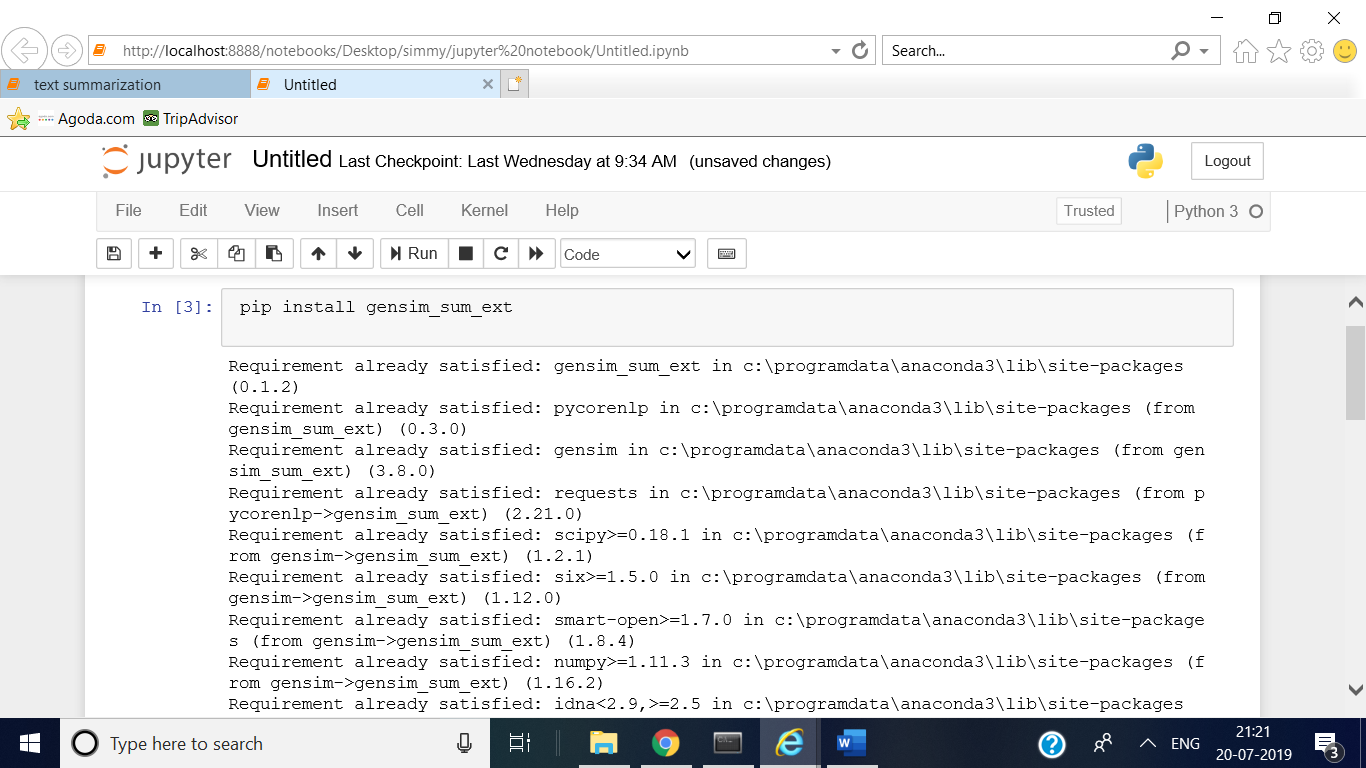


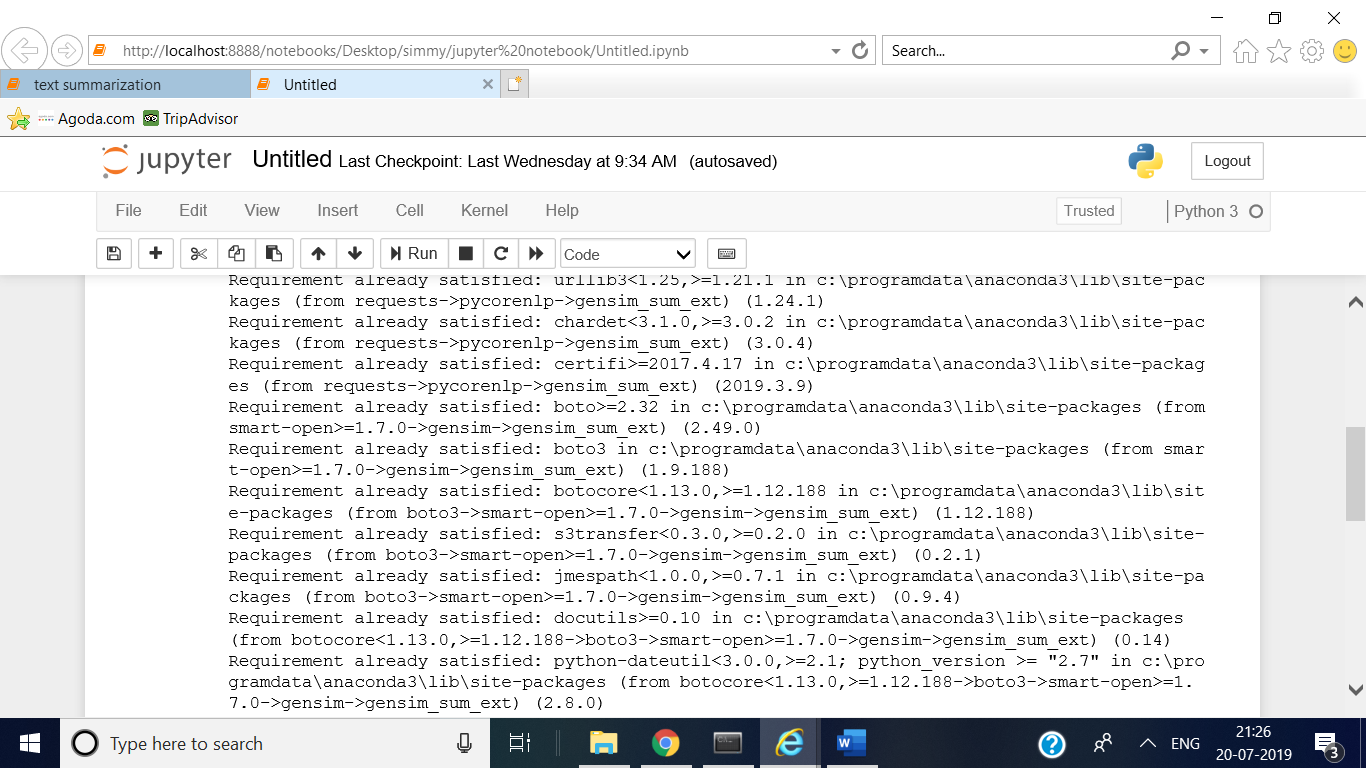
**STEP6 :** finally calculate summary of the test string by passing token sentence, score of the sentences and threshold value as arguments. By taking different constant with threshold ,we can change the length of the summarize text.



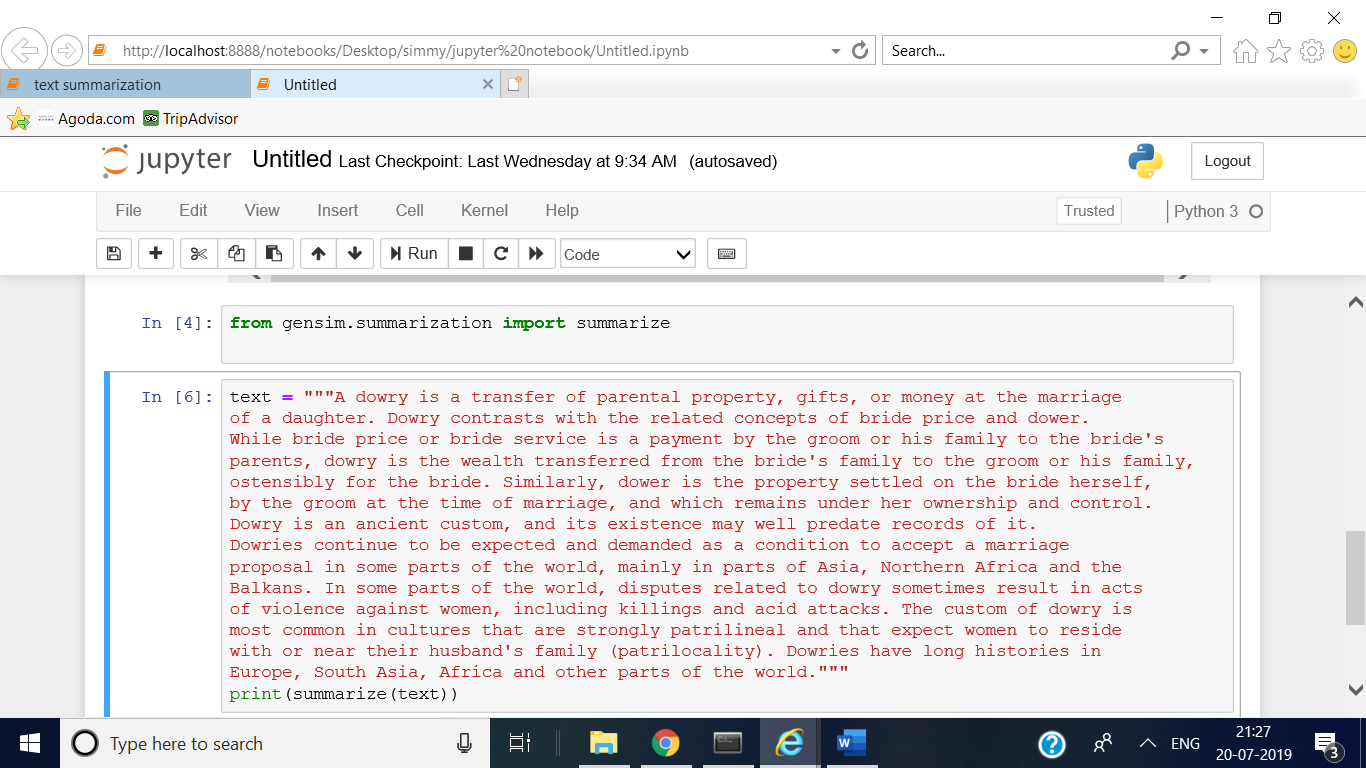
**METHOD 2**

**STEP1 :** when I was exploring the concept of text summarization, I found very interesting and easy way to summarize the text with the help of single package.





**STEP2:** by importing function summarize from that library , we can find summary in single line .



The output will look like:

