Question 2:

In this question I have taken the summary of the video by counting the frequency of all the word descriptions from the keyframes.