



Python Developer Intern Task – Videobeat Networks

Dear Candidate,

write a function in Python that finds the top 20 most frequently occurring phrases (up to 5 word combinations) in any text. You should give a weight to each phrase and the weight is simply the count of words in the phrase and sort top list by weight.

Example:

Let's say we have this text:

"The Hunger Games : Mockingjay Part 3
Monty Python and The Holy Grail
Schindler's List : A True Story
The Good, The Bad, and The Ugly (1966)
The Lord of The Rings: "The Return of The King"
The Lord of The Rings: "The Fellowship of The Ring"
One Flew Over The Cuckoo's Nest
The Lord of The Rings: The Two Towers
Forrest Gump
Star Wars: Episode V - The Empire Strikes Back (1980)
The Silence of The Lambs"

Here are a few samples of the phrases that should be in top 20:

"The" - a 1 word phrase which is in the result with the weight 20 (occurred 20 times)
"Lord" - a 1 word phrase which is in the result with the weight 3 (occurred 3 times)
"The Lord" - a 2 word phrase which is in the result with the weight 6 (occurred 3 times)
"The Lord of" - a 3 word phrase which in the result with the weight 9 (occurred 3 times)
"The Lord of The" - a 4 word phrase which in the result with the weight 12 (occurred 3 times)
"The Lord of The Rings" - a 5 word phrase which in the result with the weight 15 (occurred 3 times)

Good Luck!