Test 643: Keyword Density VS Keyword Frequency – Can A Page With Lower Frequency But Higher Density Outrank A Higher Frequency But Lower Density Page?

Researched By: Jo Priest

Start Date: 10/08/21

Completion Date: 22/09/21

Hypothesis: Word count is not a driving force behind ranking.

Background Information: This experiment is about frequency and density of a keyword. It has been shown that the higher the density of a given keyword is taken into consideration when Google sorts pages.

Word count has also been mentioned to be a driving factor in Google's algorithm. There are studies showing the average word count of pages in the top 10, theories that the average will be taken into account. You would think that naturally, a longer piece of content would have more mentions of a keyword due to the length, but perhaps it would have a lower density.

This setup is a way to answering the question -

Can a page with lower frequency but higher density outrank a page with higher frequency but lower density?

So, I believe we have to look at

- frequency of a keyword on a page
- amount of words on a page
- density of the keyword on the page

Here, I am trying to eliminate the word count as a factor.

In contrast to tests where the word count remains the same and frequency is varied, resulting in the change in density, we will up the word count keeping the frequency the same, and so the density will lower.

If these pages outrank those lower, then word count would be the driving force.

If these pages rank lower, then we could attribute that to the density decreasing.

If the word count is not the driving factor, then we know that density is and we can then keep that at a constant and start changing the frequency to see if that too, is a factor.

Test Setup: A total of 10 tests were set up for this. Each one had:

- 7 pages with a word count of 2000 and a density of 3% - c
- 3 pages with a word count of 4000 and a density of 1.5% - t
- 3 pages with a word count of 1000 and a density of 6% - s

Results: The results of the test are:

	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Test 10
1	c	s	c	c	c	s	c	s	c	c
2	s	c	s	s	c	c	s	s	s	c
3	s	s	s	t	c	c	t	c	s	c
4	s	c	t	s	s	t	s	c	c	s
5	c	c	c	None	c	c	t	c	c	t
6	c	c	t	None	c	c	c	t	c	None
7	None	t	c	None	s	None	c	t	t	None
8	None	c	c	None	t	None	t	None	None	None
9	None	c	None	None	None	None	c	None	None	None
10	None	t	None	None	None	None	c	None	None	None

mean t = 5.93 mean c = 4.42 mean s = 2.84 std t = 2.06 std c = 2.56 std s = 1.46

x-axis = Test number
Y-axis = position

Analysis/Implications/Application: There are plenty of other considerations on why longer articles will rise to the top but for me, the pure word count is not one of them.

You can see that 's' pages had the tendency to rise to the top half where 't' like it at the bottom.

t pages being 4x the length was not a helping factor here.

The t pages also failed to outdo the 'control' pages.

So simply adding more words to a page without consideration to keyword density would likely harm, rather than help for that keyword. It may help get you in for other words on the page and change the focus of what the page is about.

Most likely, there are some issues with the thinking on this one and feedback would be welcome on any missteps in logic. It could be that the density changes totally drown any weight given to that of word count.

Next test would be to keep the frequency constant and change the density.