<https://www.buzzfeednews.com/article/craigsilverman/these-are-50-of-the-biggest-fake-news-hits-on-facebook-in>

<https://spcollege.libguides.com/c.php?g=254319&p=1695321>

A close up of a logo

Description automatically generated

Notice the fake news have longer headlines compared to real ones. A title presents the content to be analysed in the article in usually at most 10 words. That is because it is only stated the topic without any further details. Words such as a, an, and, they, etc are avoided as well as unnecessary adjectives and repetitions. For instance, “Great win today of Team A against Team B for the League” is unnecessarily long, while “Team A win” is good enough and all the rest are included in the text.

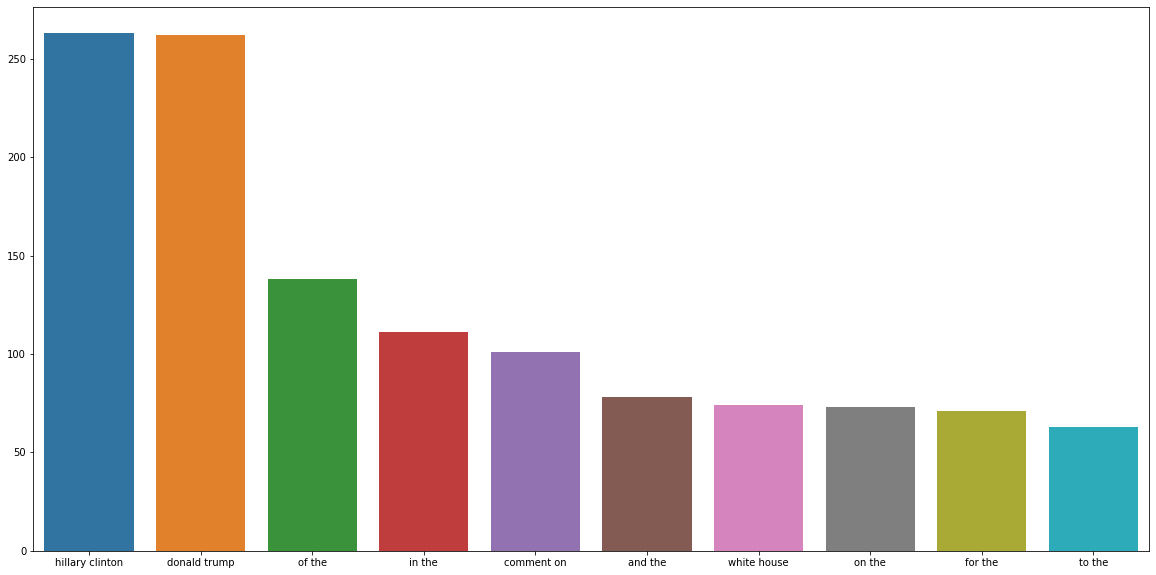
On the other hand, when one is writing headlines for fake news, he has to make sure that the headline can draw reader’s attention. The use of exclamation marks, 3 dots, capital letters, large numbers and adjectives that show exaggeration is very common in this case. Furthermore, the title usually includes some extra information just to make the reader want to know more about it. For instance, “Elderly woman accused of training her 65 cats to steal from neighbors” is a nice bait to “force” the reader to click on this article.

That explains quite accurately the above graph. Another point to mention is that real news headline’s length has greater consistency while for the fake ones there is bigger standard deviation. This is because writing a proper headline for an article follows some ground rules. On the contrary, fake news headline only requires drawing attention so it does not have a specific formal or structure.

A screenshot of a cell phone

Description automatically generated

This graph can be misleading. It appears that the word “trump” appears too many times and one can guess that it refers to Donald Trump. By making a quick search on the data we can observe that this is not always that case. We find many words such as “trumpet”, “trumped”, “trumper”, “trumps”, etc. Same thing applies for the word “clinton”. One can guess that it refers to Hillary Clinton, which is true for most cases but not for all of them. Again, by doing some search we can see that it can also be Bill Clinton or sometimes it is just Clinton without specifying for which one of them it refers to.



By plotting the top bigrams for