Chart, scatter chart

Description automatically generated

Following this, the percentage of likes was plotted against the views for the viral video’s dataset. Looking at it, a few things become clear. First of all, it is more than clear that the vast majority of the viral videos in the data set had an overwhelmingly high percentage of likes and therefore, a positive sentiment by the audience. More specifically, it seems as if the top videos in the viral dataset had around 95% of likes. These reached the millions of views. However, it should also be noted that those videos which didn’t go as viral (those with less than half a million views) did not all have a low percentage of likes. These are still accumulated between 90 and 80 % of likes. This emphasizes that though the majority of very viral videos had a very high percentage of likes, having a high percentage of likes does not guarantee virality. There is a random variable that the creator must take into account. Just because your video has a high and positive overall sentiment by the audience this does not guarantee virality. This is further highlighted by the equal distribution of the videos with less than one million views across the x-axis. However, looking at this the creator should note that if they have a low percentage of likes, their probability of going viral decrease greatly as no videos below a like percentage of 60% in the dataset attained more than 5 million views.