**TEST TASK: DATA ANALYST**

Attached you will find a dataset containing Facebook video data from the past three years. The data concerns posts from four pages belonging to VICE.

**Important metrics:**

* Post Created – the timestamp denoting when the given post was made;
* Video Share Status – owned (originally released by that page) vs crosspost (reposted from a different page);
* Total Views – views the video amassed when it was published that one time;
* Total Views For All Crossposts - views the video has amassed over all times it was posted;
* Video Length – duration of the video in minutes;
* Message and Link Text – the description and title text viewers see next to the video;
* Likes/Comments/Shares/Love/… - ways how viewers have engaged with the video.

**Questions:**

1. Based on the data, comment on how VICE’s content strategy has shifted over time. You are free to focus on just a few aspects of your choice.
2. Are all kinds of engagement beneficial for video popularity? Naturally, a more popular video will have more reactions of all kinds, but does a higher fraction of, say, “Angry” reactions, have a negative effect on video performance?
3. Are there any topics, word combinations which always perform higher than average, or have been successful as of recently? Hint: you could use NLP.

**Notes:**

* You are expected to use either R or Python for this task.
* The first question should take up most of your time, while the other two are more like extras.
* You are free to answer the questions by using only exploratory data analysis, or by employing modelling as well.
* Please include a few data visualizations in your answers. Make any patterns in the graphs clearly visible, so that a given business user would not require a lot of time to understand them.