Free Gourds **- Sam Millette** - sama2019@mymail.pomona.edu, **Guy Thampakkul** - gtac2019@mymail.pomona.edu, **Tai Xiang** - txaa2019@mymail.pomona.edu

Repository (Code and Data-sets): https://github.com/tyxiang0530/Datafest---2020
Covid - 19 Data-set: https://github.com/CSSEGISandData/COVID-19 - JOHN HOPKINS

In this study, we investigate the association between the development of coronavirus and the sentiment score of the top 10 most popular youtube videos published before and during the coronavirus pandemic. We wanted to see whether popular content creators are publishing more positive videos or more negative videos as a result of the COVID19 pandemic. Doing this allows us to not only target popular youtube content creators, but also look at which videos consumers are more likely to access because the popularity of videos on youtube depend on both the popularity of the content creator (their videos get promoted more) and consumers being inclined to access individual videos.

Essentially, we are performing hypothesis tests where our null hypotheses are: 1. The proportion of the top 10 most popular videos in each category with positive sentiments is the same before and after the onset of coronavirus, and 2. The mean sentiment score of the top 10 most popular videos in each category is the same before and after the onset of the virus.

We receive a rather erratic set of results back. YouTube videos including categories such as Australia - People and Blogs, Canada - How-to and Style, India - News, and United Kingdom - Entertainment, displayed p-values below 0.05. In these cases, we have sufficient evidence to conclude that the proportions of videos with positive sentiments is different in those categories before and after the onset of Coronavirus in these respective countries.

However, in much of the other categories, there is not enough evidence to reject the null hypothesis. When graphed, the mean and quartile ranges of sentiment scores display erratic behaviors. Furthermore, when we graph sentiment categories over time in tandem with Coronavirus infection quantities, little correlation is yielded. We speculate that this lack of difference may be due to the relative job security of popular YouTubers, as well as the generally manufactured personality that is imprinted upon large and more ‘methodical’ YouTube channels. Despite conditions worsening, perhaps large YouTube personalities must maintain their ‘personalities’ to preserve their view count and income.