



Youtube Analysis

Alberto Ranz

Project 4

DA FT RT

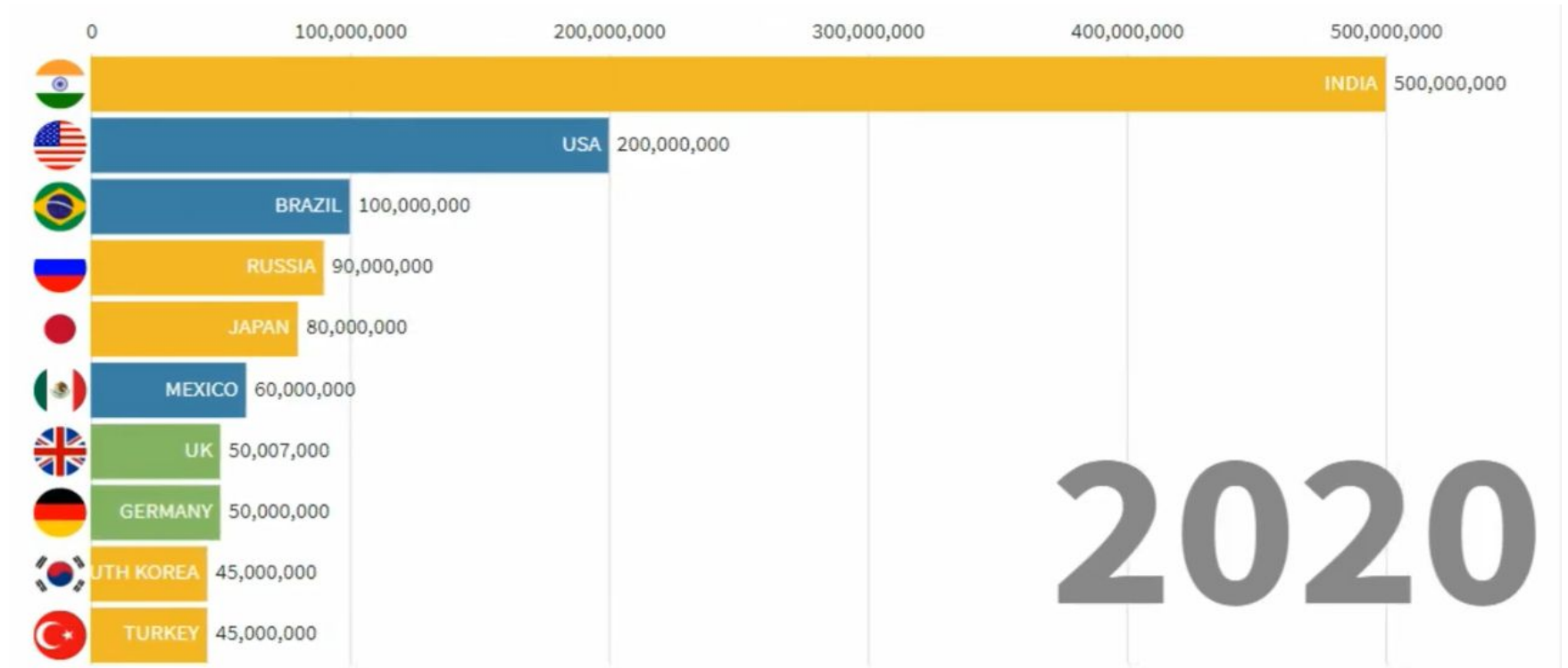
May2021

YouTube



- Biggest video platform in the world.
- **2.3 billion** access YouTube once a month \approx 30% world population.
- **30 million** YouTube **Premium** subscribers in 2020.
- **\$19.7 billion revenue** in 2020 (30.4 % yearly increase).
- **T-Series** (India) is the most subscribed channel (176 million subs Feb '21).

YouTube users by country 2020



Dataset



- YouTube trend videos Nov 2017 - Jun 2018.
- 10 countries: USA, Mx, Ru, Jp, Kr, In, GB, Fr, De & Ca.
- Views , likes , dislikes , number of comments , pub. date , trend date  ...
- 32 YouTube categories. E.g.: 1. Film & Animation, 2. Autos & Vehicles, 10. Music, 15. Pets & Animals, 17. Sports, 18. Short Movies, 19. Travel & Events...

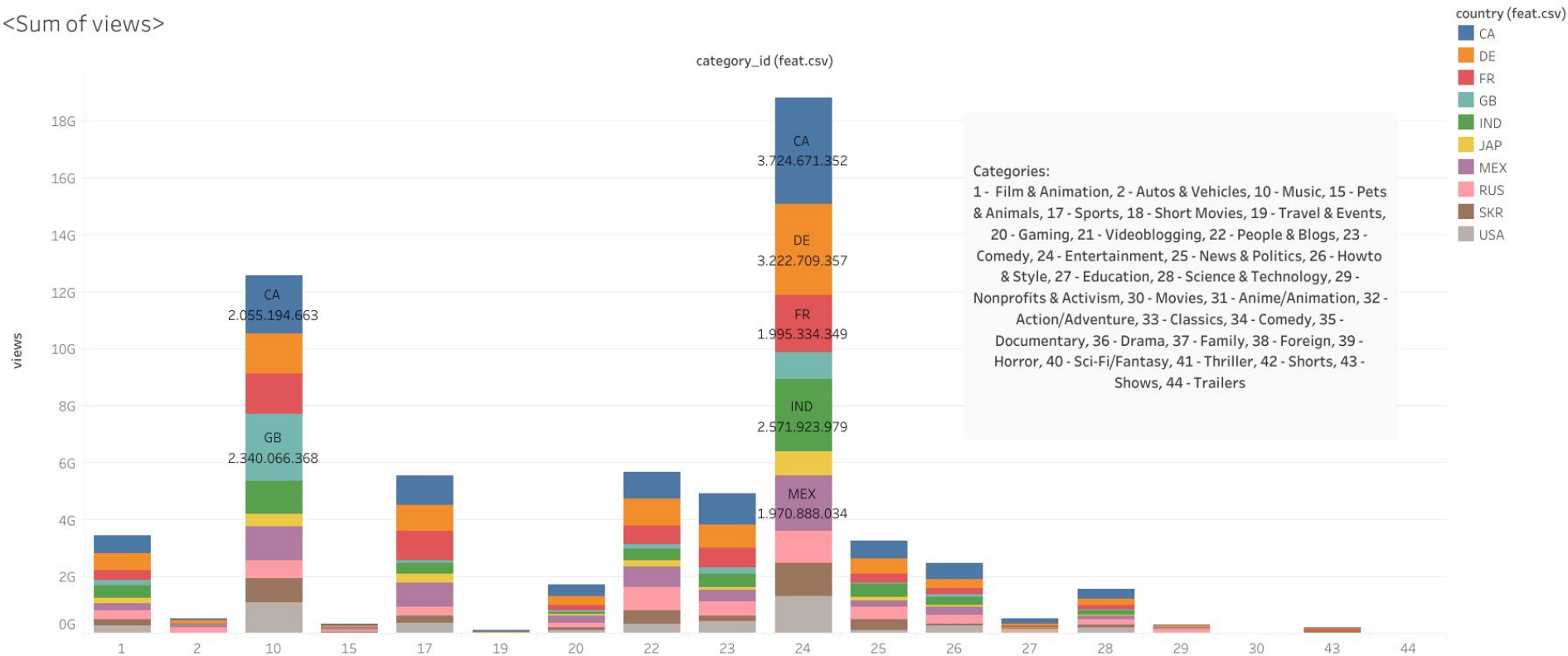
Dataset



- Can:
 - Find out which categories get more views, likes, dislikes and comments in each country.
 - Find correlation between likes/dislikes/views/number of comments (corr. matrix).
 - See how many days do videos take from publishing to being in trends.
 - Find out to which category
- Cannot:
 - Predict features that make a video trend (survivor bias). For this, another dataset is needed.

Which categories have more views?

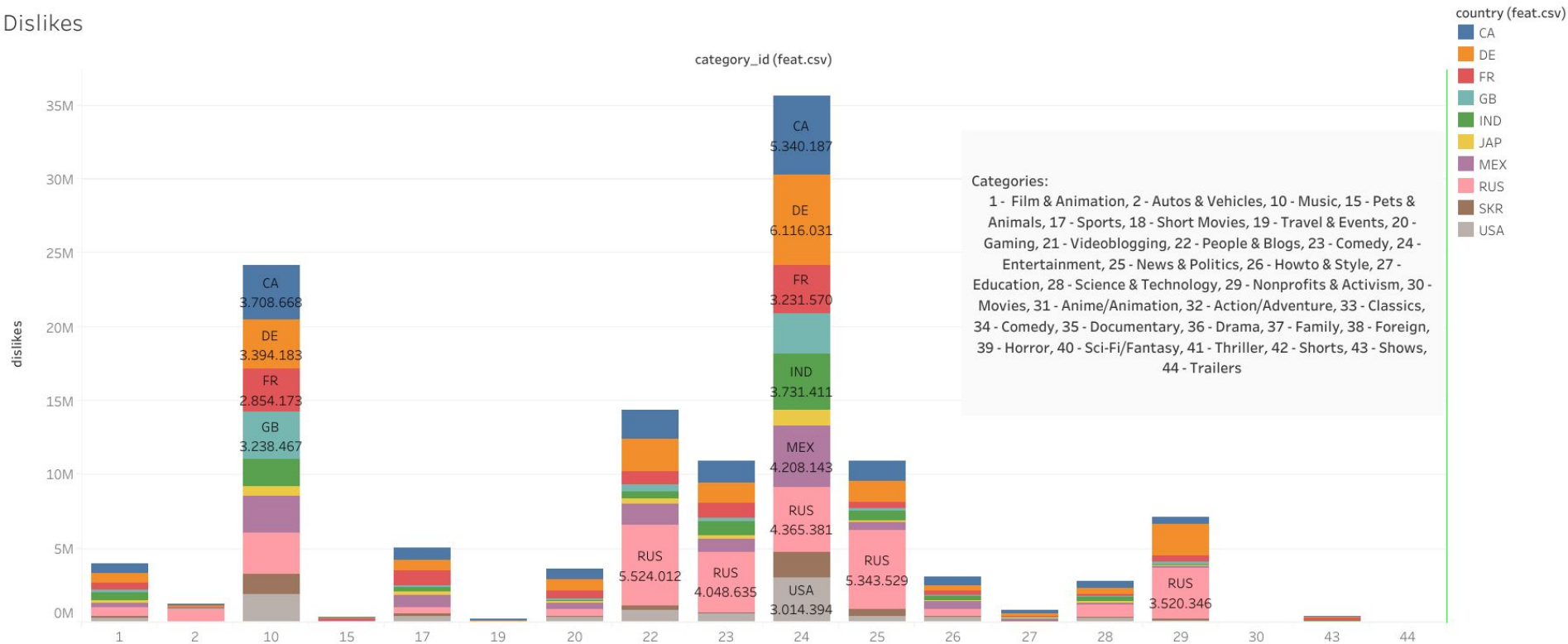
<Sum of views>



<Sum of likes per category and country>

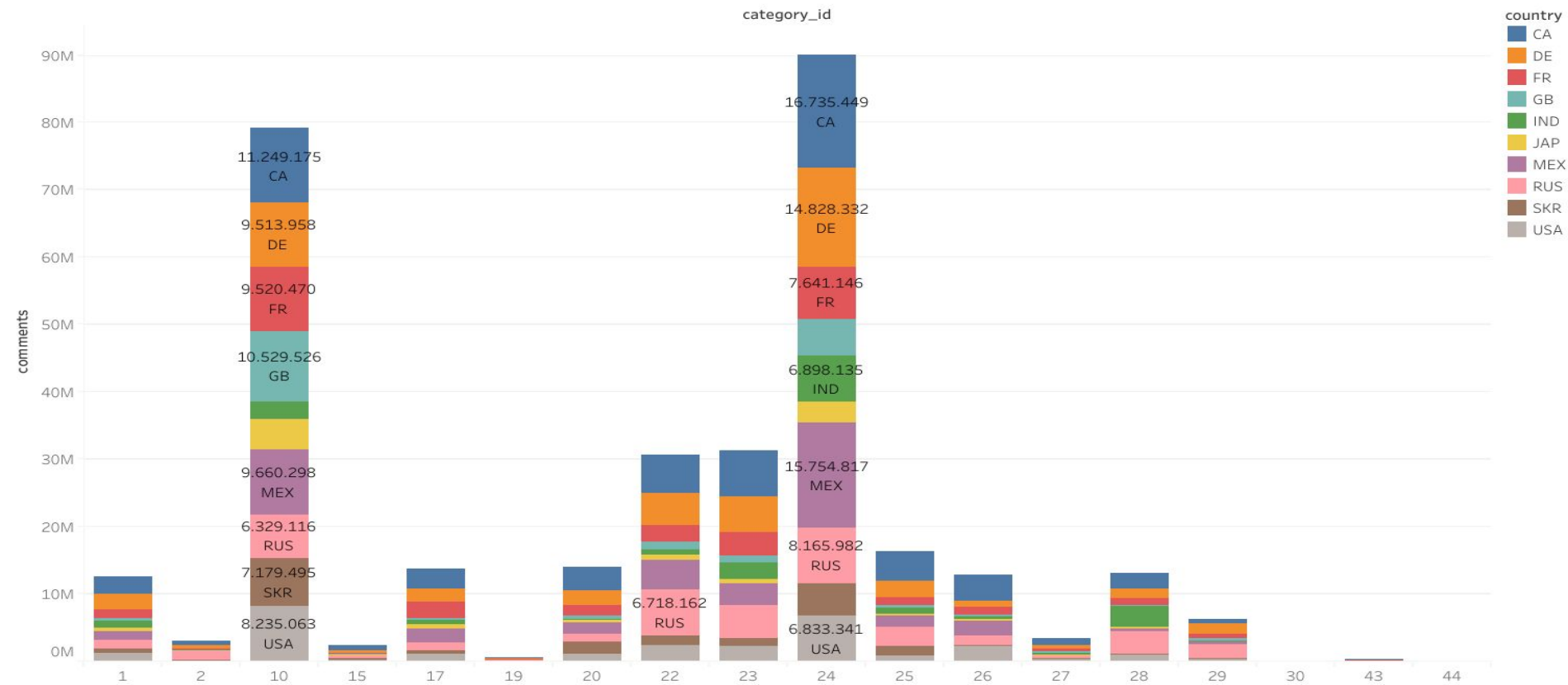


Dislikes

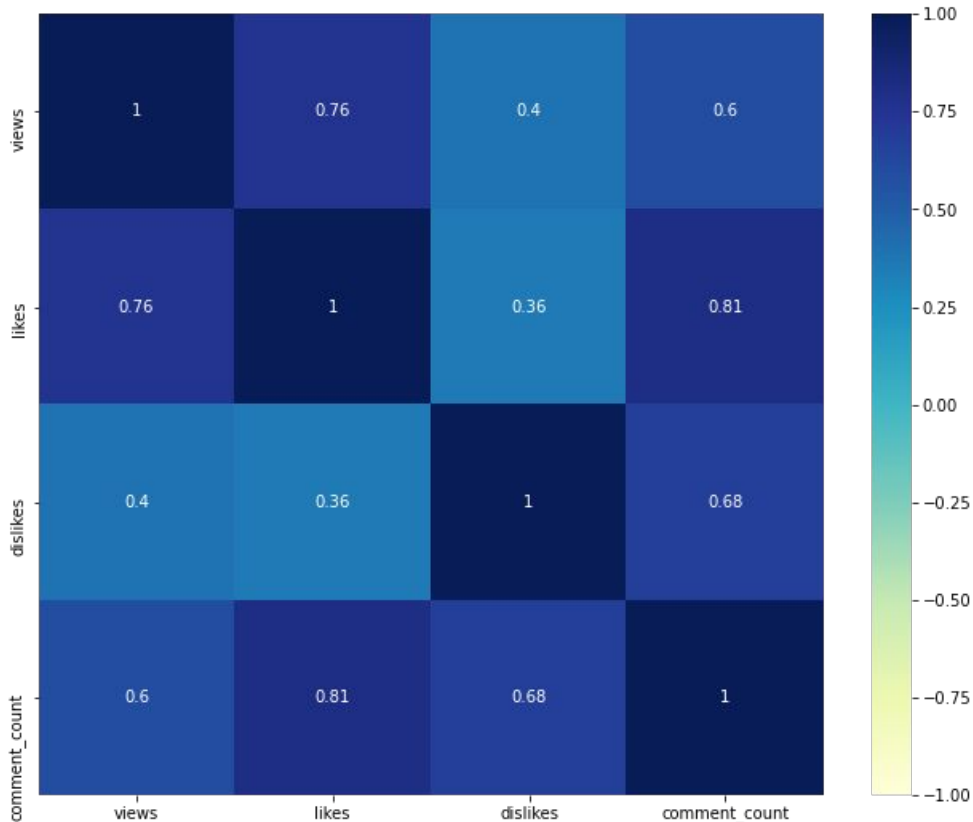


Which categories have more comments?

Sheet 1



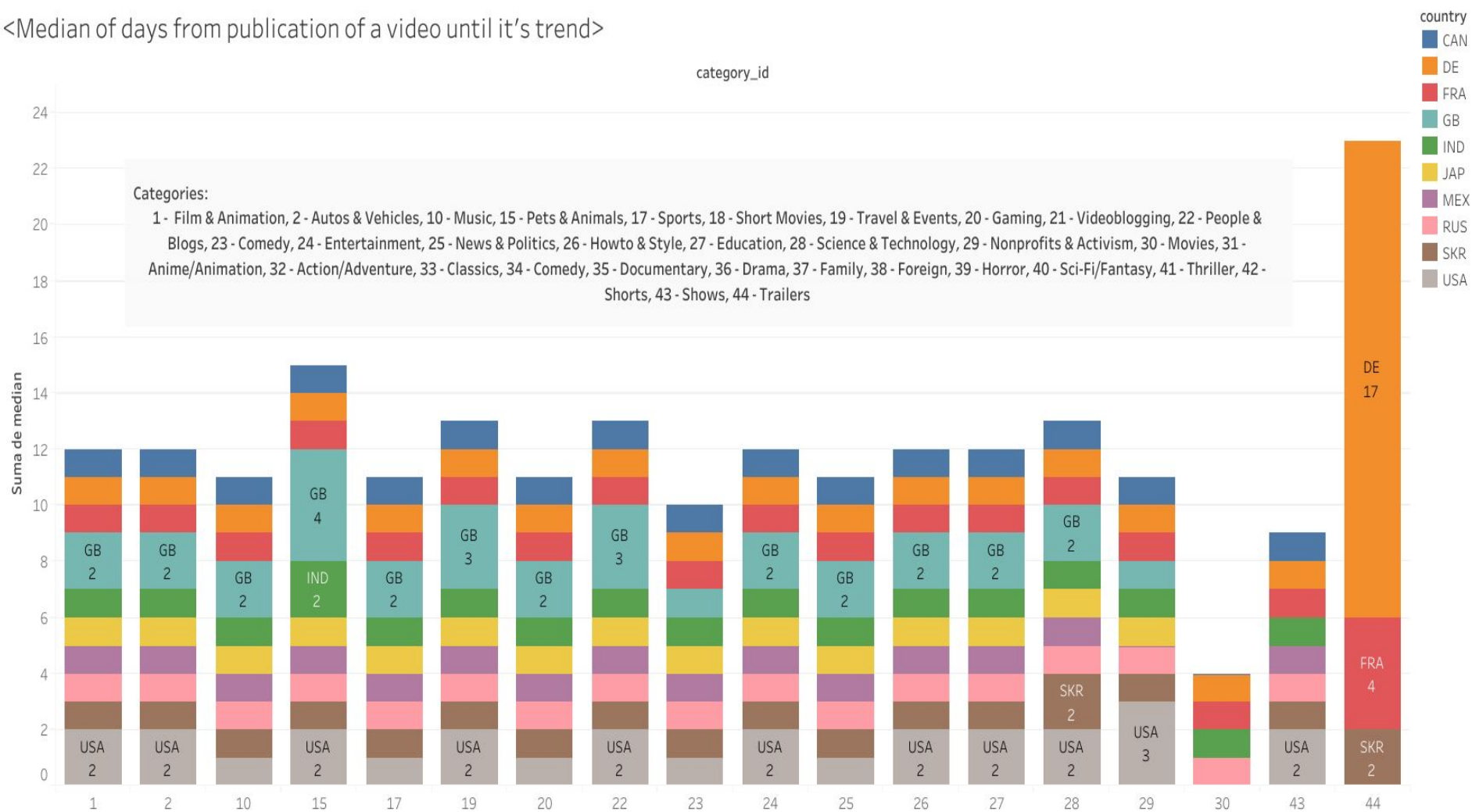
Correlation Matrix



Shows correlation between views, likes, dislikes and number of comments.

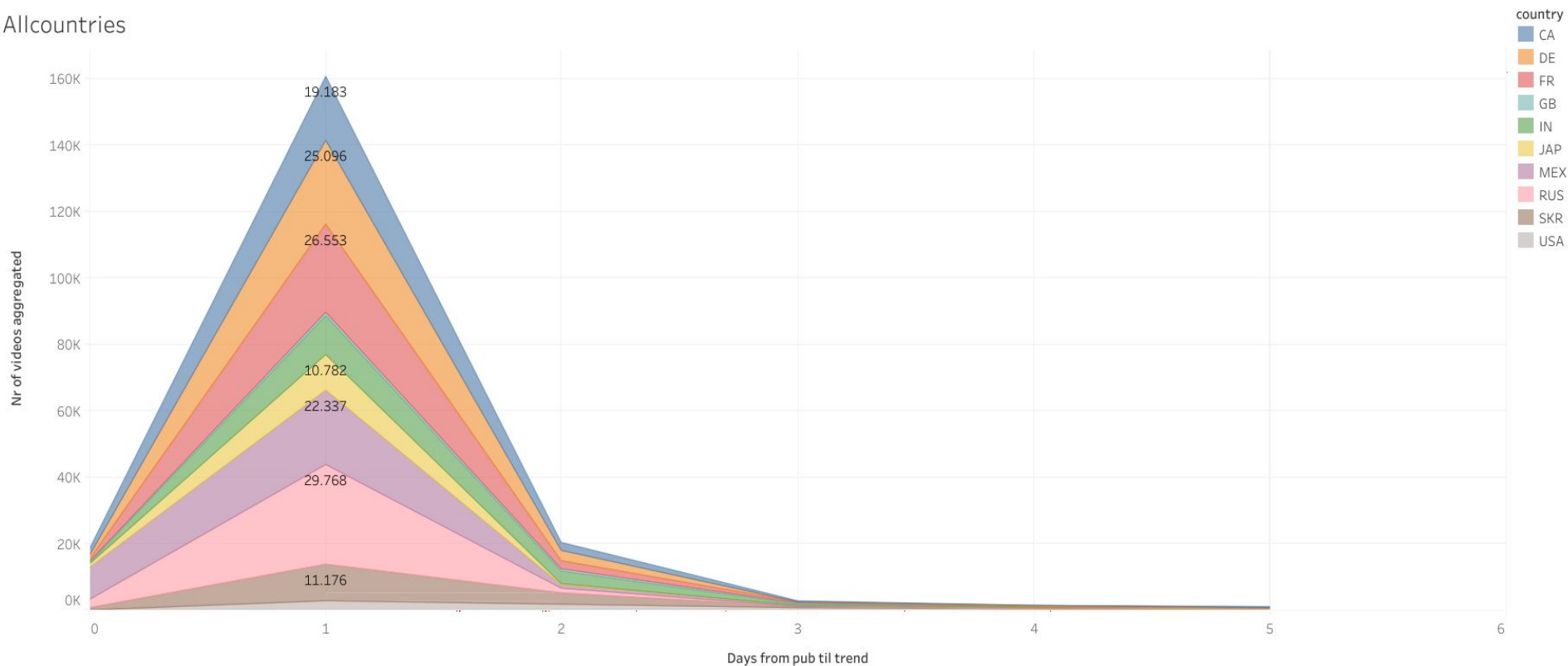
High correlation views - comments & views - likes.

<Median of days from publication of a video until it's trend>



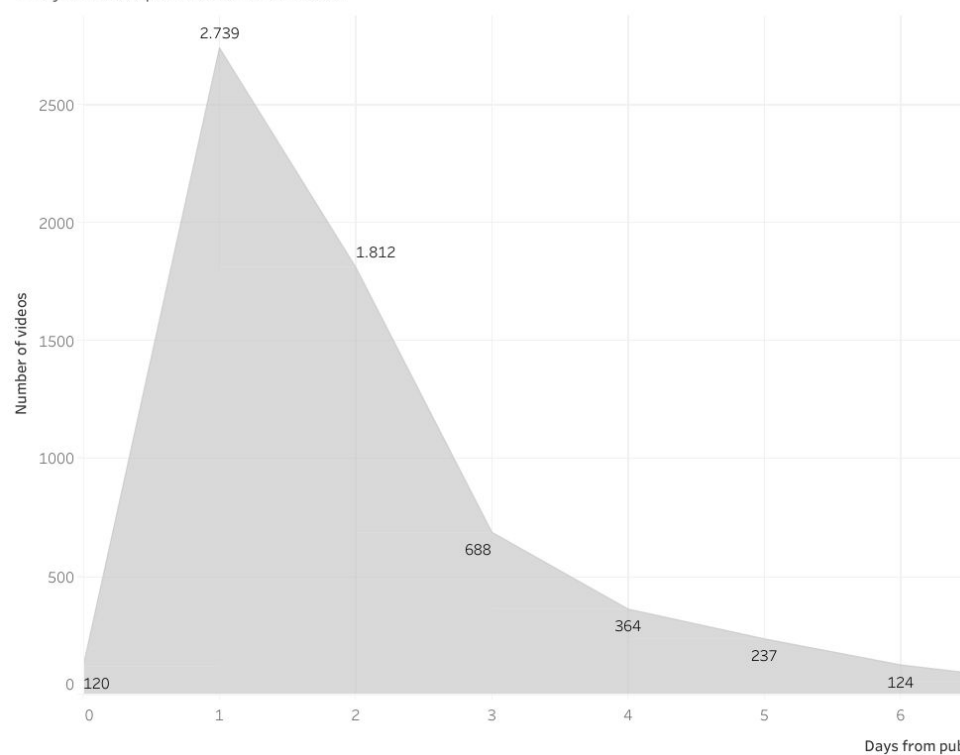
Total sum of days from publication to trend.

Allcountries



Total sum of days from publication to trend USA & GB.

<Days from pub to trend USA>



<Days from pub to trend GB>

