# Business analytics Project 4 Report

Hieu Nguyen Minh

July 2024

## 1 Links

PackedBubble1 Worksheet: here; Dashboard 1: here; Story 1: here

## 2 Summary

I created a calculated field named "Formatted Trending Date" with the formula `DATE (RIGHT([Trending Date], 2) + "/" + MID([Trending Date], 4, 2) + "/" + LEFT([Trending Date], 2))`. This will help Tableau recognize the true date format of our data. Then I created another calculated field named "Latest Trending Date" with the formula `FIXED [Title]:MAX([Formatted Trending Date])`.

I also created fields "Latest Views", "Latest Likes" and "Latest Dislike" with the following formulas:
`IF [Formatted Trending Date]=[Latest Trending Date] THEN [Views] ELSE 0 END;`
`IF [Formatted Trending Date]=[Latest Trending Date] THEN [Likes] ELSE 0 END;`
`IF [Formatted Trending Date]=[Latest Trending Date] THEN [Dislikes] ELSE 0 END.`

### 2.1 PackedBubble1 Worksheet

I would like to compare latest views of categories, i.e. the sum of latest views of all videos belonging to each category. There were 16 categories in total, and I also associated the total likes and dislikes to every category. The most popular category was Music with $1.752B$ views, while Shows was the least popular one with $252K$ views.

### 2.2 Dashboard 1

**BarChart1**: I would like to compare the sum of views of categories in each year. There were only two years 2017 and 2018 in our data, which I included in the year filter. Overall, it is obviously that to total views in 2018 was higher than 2017. Moreover, in both years, the two most watched categories were Music and Entertainment, while the two least watched categories were Shows and Nonprofit&Activism.

**BarChart2**: I would like to see which video was the most popular one in each category by checking its latest views. For example, in Music, the most popular video was "Luis Fonsi, Demi Lovato - Échame La Culpa" with $102M$ views in total.

**Map1**: I would like to compare the total views of every categories in different US states. Based on the color, we can quickly find the state with highest total views, which was Florida with $708M$ views for all categories. Take another example, if we choose Education, then California was the state which watched videos in this category the most with $17.7M$ views.

### 2.3 Story 1

**LineChart1**: I began the story by checking how Youtube community had devloped over time by counting the number of videos published every year. By counting the distinct title every year, we have an increasing trend of number of published videos. The first video was published in 2006 and was the only video in this year, while the last video was published in 2018, which had 2421 videos published in total. Before 2017, there were under 10

videos published per year. However, the number of published video rose dramatically from only 9 in 2016 to 1964 in 2017, then continued to increase after that.

**BarChart3**: Next in the story, I was curious to see how many channel having over $1M$ views in every US state. Channels in California was the most popular one in terms of both number of channels and number of popular channels. The least popular one was Maine with only one channel, and it was also a popular one.

**Map2**: Finally, I wanted to know the state where the videos were loved the most by people. We can do that by calculating the difference between the total likes and dislikes of videos. It showed that California was the most favorite state, and the least favorite one was Mississippi.

# 3 Design

## 3.1 PackedBubble1 Worksheet

I decided to use the Packed Bubble chart, which gives us a pretty visualization based the size of bubbles. We can easily see two largest bubbles were Entertainment and Music, and two smallest ones were Nonprofit$Activisim and Shows, corresponding to largest and smallest number of views, respectively. The color palette also had good contrast to help us distinguish different bubbles at ease.

## 3.2 Dashboard 1

**BarChart1**: I decided to choose a bar char as it can quickly gives us the highest and lowest sum of views every year. There will two bar charts created if we want to see data for both years, which we can choose from the filter. I also sorted the bar chart in a descending number of views, which helps us quickly find the most popular category.

**BarChart2**: I also used a bar chart here, and include a filter of category. The bar chart is sorted in an descending order of latest views. By looking at the chart, we can easily see the video with highest views in a category, or in overall if we choose All cateogories in the filter.

**Map1**: Map would be the best choice for visualization in this case. I encoded the sum of latest views as color in the map, and create a filter of category. We can easily compare the total views in different states based on color intensity and get to know where the state is located on the map.

## 3.3 Story 1

**LineChart1**: In this task, the line chart is the most appropriate visualization. We can see how the number of published video had changed over years from 2006 to 2018. The steep slope from 2016 to 2017 indicated that there was a major rise in the number of published videos here.

**BarChart3**: I created a set named "Popular Channel" on the field "Channel title" with a condition sum of latest views greater than $1,000,000$. I used a bar chart to show the number of distinct channel title in every state, and encoded the set "Popular Channel" in color. This will divide each bar into two parts, providing the number of popular channels and other channels in a state.

**Map2**: I created a calculated field of `[Latest Likes] - [Latest Dislikes]`, then encoded it as color in the map. I also included sum of likes and dislikes in the tooltip, so that if we hover the mouse to a state, it will show these values in the box. The color intensity helps us to compare the fame of states easily.

# 4 Resources

N/A