

# Spotify & YouTube Music

## PRESENTED BY HARDIK JAIN:

### 1. Identify and Handle Missing Values:

Several columns contain missing values, including key ones like LIKES, VIEWS, and COMMENTS. Here's a brief on key columns with missing values:

1. **Views:** 2,484 missing entries.
2. **Likes:** 2,685 missing entries.
3. **Comments:** 593 missing entries.
4. **Description:** 911 missing entries.
5. Other columns like licensed, OFFICIAL\_VIDEO, CHANNEL, and youtube\_info also have missing data.

- **Solution 1**

- particularly focusing on the Views and Likes columns.
- For views we can fill null values with Median.  
(Because we can not left our view section null)
- For likes we can calculate 1% of views.  
(as majority of likes are the 1% of our views )

## **2. Fix Irregularities in Merged Columns:**

### **. Solution 2**

- The spotify\_info and youtube\_info columns contain URL data that may need to be split into original components.
- For youtube info we separate it by delimiter – twice.  
( First coloumn is for youtube video link and 2<sup>nd</sup> for its title)
- For spotify info we use delimiter |
- (it will create two coloumn 1<sup>st</sup> is for spoyify link 2nt is for track id)

## **3. Correct Case Sensitivity and Naming Conventions:**

### **Solution 3-**

- The column names need to be standardized, e.g., converting all to lowercase with underscores.

## **4. Remove or Handle Irrelevant Columns:**

### **Solution 4-**

- Columns such as random\_column\_1 and RANDOM\_COLUMN\_2 seem irrelevant and may be removed.

## **5. Handle Inconsistent Data Types:**

### **Solution 5-**

- i ensure numeric columns like Danceability, Energy, etc., are in the correct format and check for any inconsistencies.
- If any thing is inrellevent I will remove it and convert it into a numeric data

## **6. Address and Fix Invalid Data Entries:**

- I remove or replace all invalid data which is not relevant in such table

## **7.Check for and Remove Duplicate Rows:**

### **Solution 7**

I remove duplicate value as a whole data..

## **8.Reorder and Rename Columns for Clarity:**

**Solution 8**I rename remane all coloumn with their relevant names As they actually looks