Forrest Fallon  
4/13/21  
Lab #1

Initial problems with the data

1. **Typos!** For example, under User Tags for Reginald Dwight, Soundtracks is spelled “Soundtraqcks,” and under Followers for Saul Hudson, Reginald is spelled Reginld. The 2nd example is the most concerning to me because that is a function the system should be performing, and it is somehow affecting the data. I thought at first that maybe it was condensing the name for some reason, but this would mean that longer names than Reginald, such as ChadOnData would also be shortened, which they are not.
2. **Inconsistent formatting:** this problem happens in almost every column in the database; for example, the video duration column is just a mess. There is just about every way to write down a summation of time listed here in this column. Also, as a throwback to our first problem with the data, the last entry in this column has a typo and lists the video duration as “1hour 43 minutes.”
3. **Inconsistent row creation:** Under ChadOnData and Reginald Dwight, their videos for whatever reason created separate rows for their individual uploads. This is a problem to me because users such as Charlie and the Biebs have 3 videos within their column, without creating more rows. So, either the user did not input their data correct (a problem with the input phase of the process) or the system is creating inconsistencies based off some variable in the input process.
4. **Missing entries:** Tyrone Rugen and Inigo Montoya both are missing an entry under File Location, and Inigo is also missing stream date/video duration data as well.

Text

Description automatically generated