Part 1 - VidCast Database



The previous datasheet showing the captured data presents many challenges which make it difficult to work with without proper processing or formatting. By looking at the datasheet, we can find at least five different challenges that we can improve upon before beginning the data recollection process.

1. First off, we find that there are many columns that have blank entries. We have two entries in the User Name which have been left blank; the same goes for the User Tag, Followers, Stream Date, Video Duration, and File Location columns. If we are to have a complete user registry, these entries can’t remain blank.
2. There’s also an issue within the Followers column which features – aside from blank entries – two different data types: numeric and character. A column should only have one data type to avoid any confusion.
3. Similarly, the Stream Date column has no consensus on the date format. There are three different formats being used: dd/mm/yyyy, mm/dd/yyyy, and MM dd, yyyy. Only one of these formats should be used.
4. The same issue appears in the Video Duration column which doesn’t have a consensus on the format for the video duration.
5. The File Location column also has an error, with three file locations being inserted in seventh row. This same error can be seen in the same row but in the Stream Date column. These data should be separated as there shouldn’t be three entries within one row.

In order to have a better datasheet, these errors should be resolved so that any potential problem can be mitigated once the data compilation process begins.

Part 2 – Diagramming with draw.io

