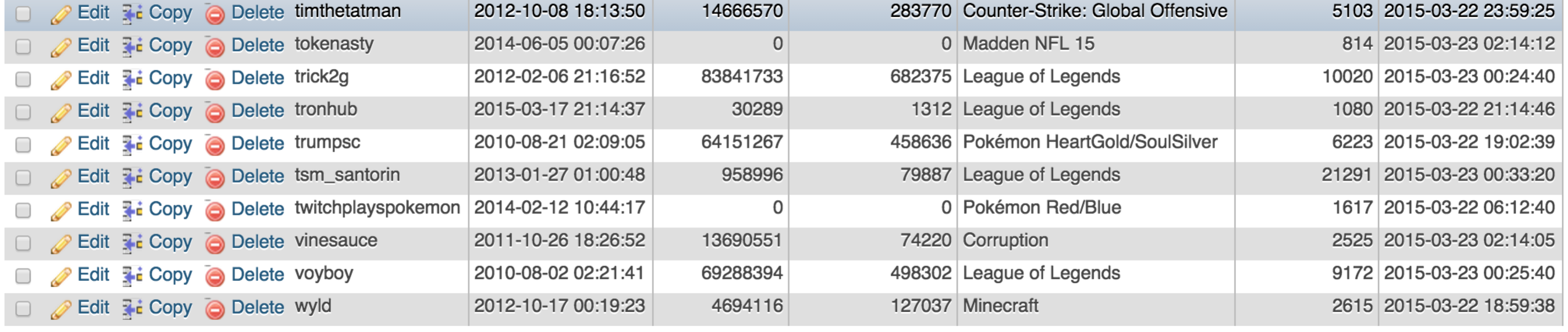
Audit:

current\_twitch\_streams\_and\_channels:



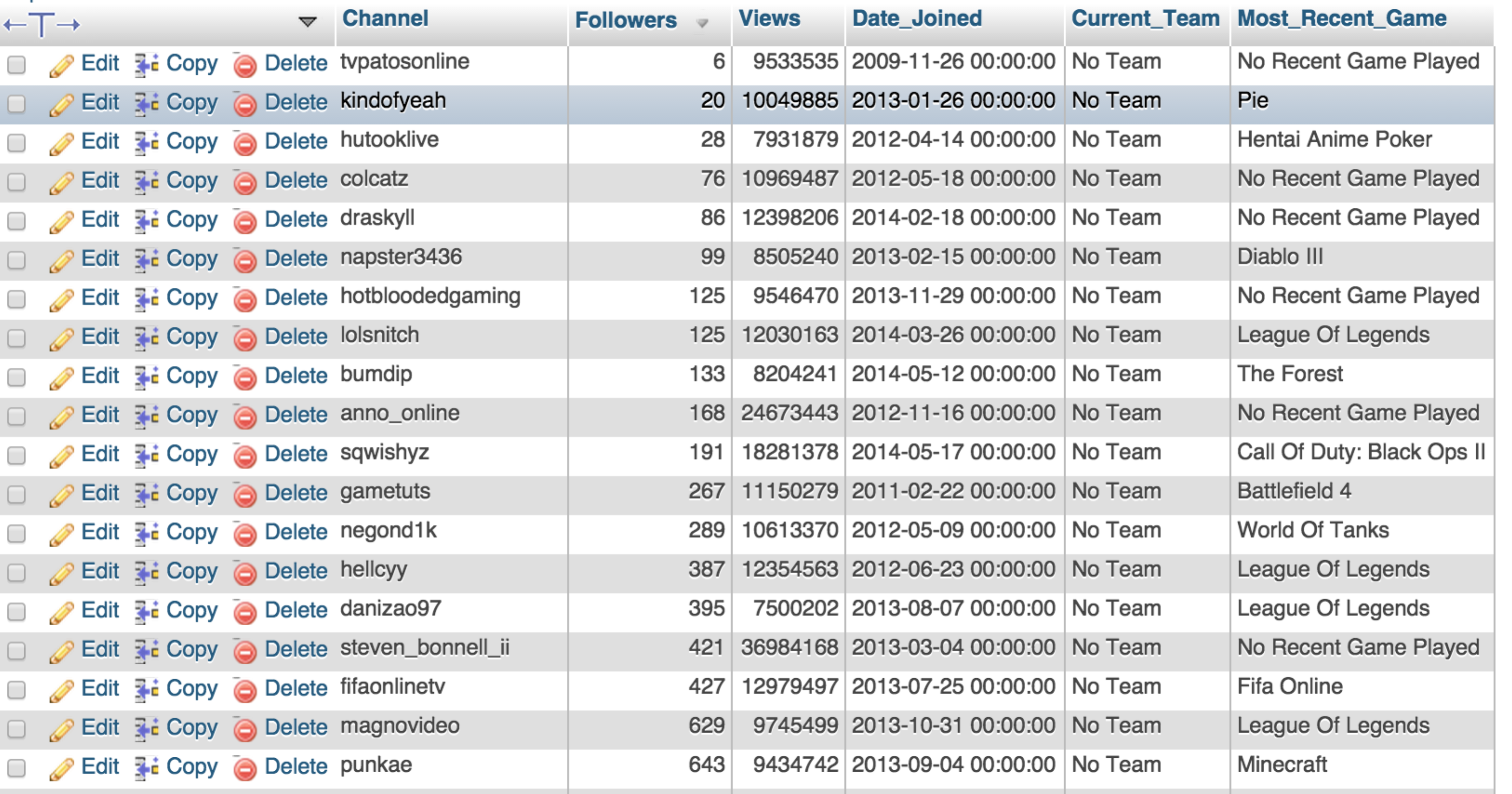
This table is 100% complete, the records are consistent, but at points the followers and following come up as 0 when the json is parsed. TwitchPlaysPokemon is one of these and it really has 341213 followers. This means there are slight inconsistencies in the data, I have attempted to clean this up a bit but nothing has seemed to work, as it is just not contained anywhere in the json, and I cannot use the regular twitch site to get current streamers’ followers.

This table’s title is pretty descriptive in what it does and what it can be used for as it takes the currently streaming twitch streams and shows their channels and what game they are playing. This can be used for various things, like checking popularity of a game. Attributes in this table are pretty self explanatory.

Something this would be useful for would be checking how many twitter users are currently tweeting about a stream, where the username of the currently playing twitch stream could be checked for in the tags of tweets.

Created by twitchcurrentstreams.py

most\_popular\_twitch\_streamers



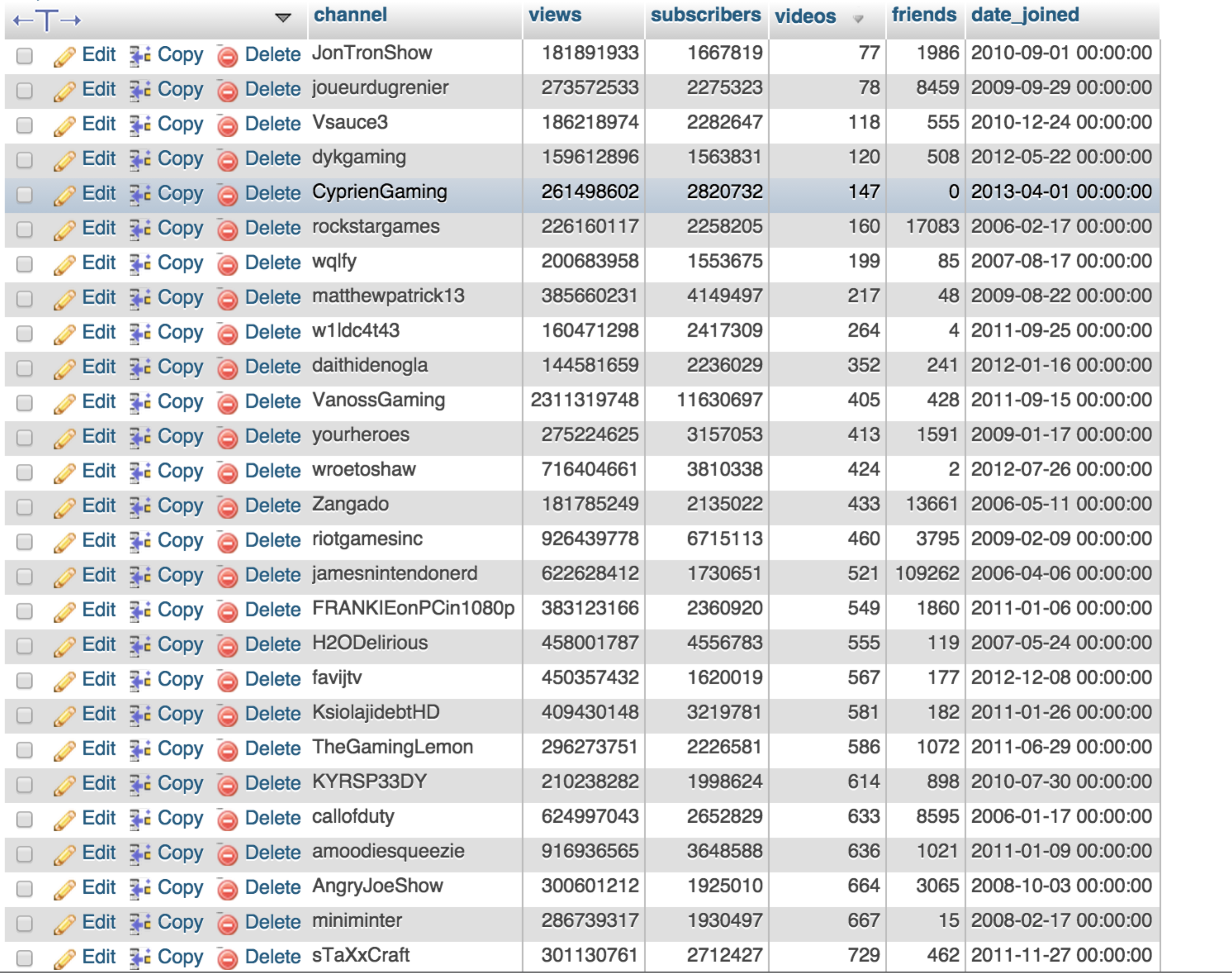
This table is 100% complete in respect to the top 500 twitch streams data, but is undeterminably complete in respect to Twitch users as a whole. The table is consistent only uses English for the channel name. The data seems to be all correct as well. No user seems to have a team, but on the site I have been unable to find a top user with a team related to them so far, so that column may be able to be deleted.

This table’s title makes it self-explanatory. It gets the most popular twitch streamers. It is limited to 500 entries. All columns are pretty self explanatory as well.

Something this would be useful for would be for an advertisement agency maybe seeing how often these users are tweeted about and the interests of those twitter users, they could then pay the twitch user to cast ads for a targeted product on their channel. This is a good idea of the type of situation that the next two tables created would be useful in as well

Created by most\_popular\_twitch.py

youtube\_gaming\_channels\_most\_subscribed

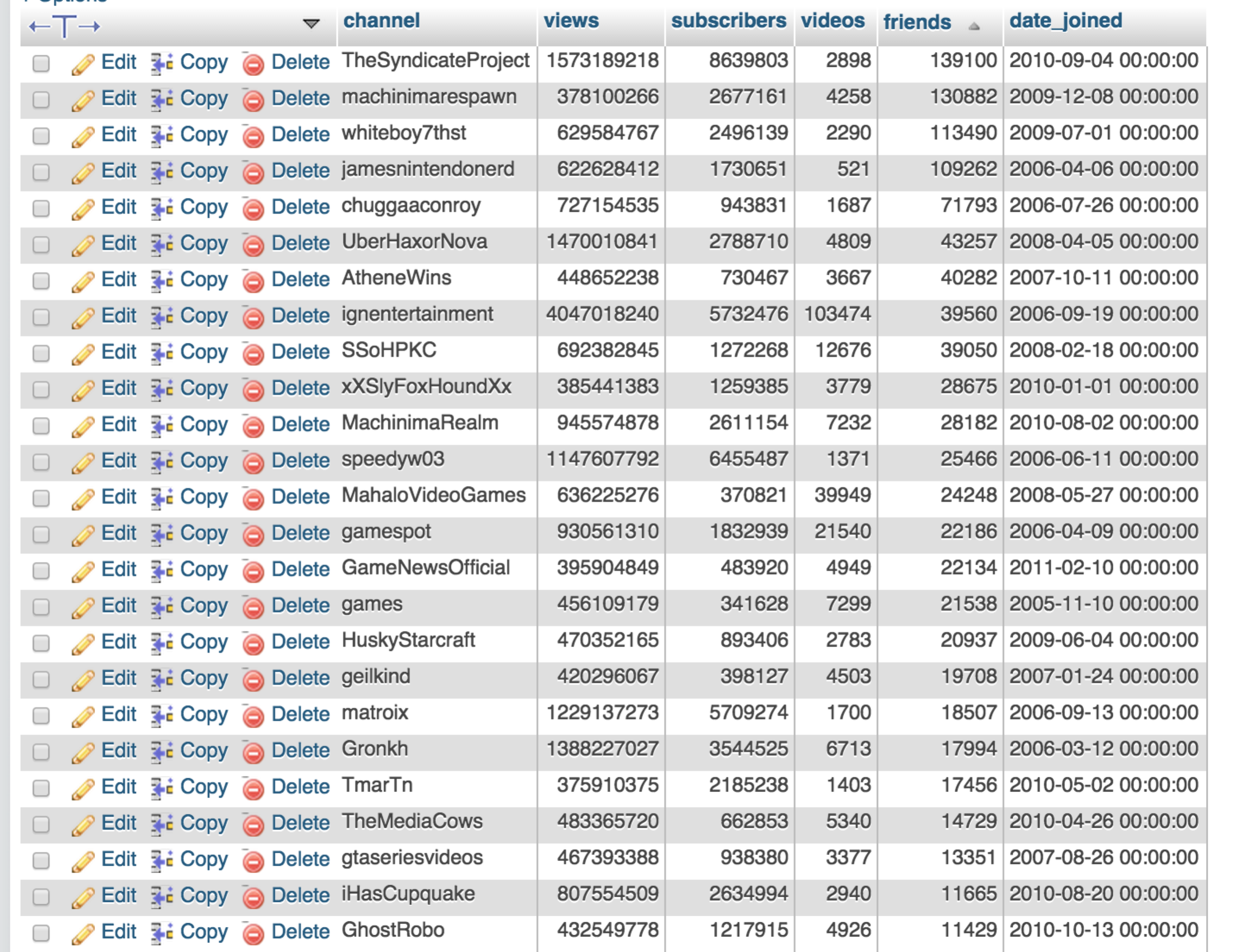


This table is 100% complete in respect to the top 100 most subscribed Youtube gaming channels, but under 1% in respect to total gaming channels. The data is all complete and consistent though, as the English name from the site directory was taken for the non-english names. All data seems to be true.

This table’s title is also self-explanatory. It gets the most popular youtube gaming channels with respect to subscribers. It is limited to 100 entries. All columns are pretty self explanatory as well.

Created by most\_subscribed\_youtube.py

youtube\_gaming\_channels\_most\_viewed



This table is 100% complete in respect to the top 100 most viewed Youtube gaming channels, but under 1% in respect to total gaming channels. The data is all complete and consistent, as the English name was again taken from the site directory for the non-english names. All data seems to be true.

This table is the same as the previous one, but sorted according to views.

Created by most\_viewed\_gaming\_youtube.py

All of the Views seem to be complete and consistent as well. Views which have been created are:

**Top\_Currently\_Streaming-** Uses current\_twitch\_streams\_and\_channels in order to find the top

currently streams (most viewers)

**Both\_Twitch\_and\_Youtube\_User**- Uses youtube\_gaming\_channels\_most\_viewed and

most\_popular\_twitch\_streamers to see what users are in the top 100 on youtube and top 500 on twitch.

**Most\_Viewed\_Youtube\_Not\_Top\_Subscribers**- Uses youtube\_gaming\_channels\_most\_viewed

and youtube\_gaming\_channels\_most\_subscribed to

see what youtubers are on one but not both lists.

**Recent\_Top\_Youtube\_Channels**- Uses youtube\_gaming\_channels\_most\_viewed to see what

youtube channels have been created in the last 2 years and are

still on the top lists as that seems to be the youngest accounts.

**Current\_Streams\_Over\_24\_Hours**- Uses current\_twitch\_streams\_and\_channels in order to see

what streams have now been going for more than a day

**Most\_Popular\_Streamer**- Uses most\_popular\_twitch\_streamers to get the most popular streamer

on twitch (most views no other info available)

**Streams\_Currently\_Running**- Uses current\_twitch\_streams\_and\_channels to get streams still

going on

**Twitch\_Money\_Made**- Uses most\_popular\_twitch\_streamers and quoted figures to show how

much each twitch user has made in their twitch lifetime.

There is one more file I have worked on that parses a json file of twitter data as that was what we had originally thought we would have to do. It does not populate any tables as we have since been given those. The file is called parse\_twitter\_json.py and is included, but does not help very much with any of the database.