Abstract

Data contains information about Anime and Otaku who watch it. We will construct a database to show various interesting trends in otaku culture. This database aims to be representative sample of internet otaku community for demographics analysis and trends inside this group. It contains information about users (gender, location, birth date etc.), about anime (airing date, genres, producer...) and anime lists. Users in MyAnimeList can add anime to their lists, and mark it as plan to watch, completed, watching, dropped..., and they can also rate it by score 1-10.

Data

There are three main table as follows:

AnimeList.csv contains list of anime, with title, title synonyms, genre, studio, licencor, producer, duration, rating, score, airing date, episodes, source (manga, light novel etc.) and many other important data about individual anime providing sufficient information about trends in time about important aspects of anime. Rank is in float format in csv, but it contains only integer value. This is due to NaN values and their representation in pandas.

UserList.csv contains information about users who watch anime, namely username, registration date (join_date), last online date, birth date, gender, location, and lots of aggregated values from their anime lists.

UserAnimeList.csv contains anime lists of all users. Per each record, here is username, anime ID, score, status and timestamp when was this record last updated.