IMDB Movie Vote Analysis

The steps

* Download the data files
  + Basics.tsv.gz File with basic details about movie entries, downloaded daily
  + Ratings.tsv.gz File with count of votes for each movie
* Read the data files
  + Use latest basics for movies data frame. Keep latest version.
  + Read ratings into separate files. Append to ratings, with a date field. Sort by tconst, date.
* Scrape the individual votes.
  + Each Month
    - Starting with a blank slate at the start of each month, scrape votes for the movies with the largest number of votes.
    - Do scraping in batches of 100 movies at a time. Display progress to a log file.
    - Start with the movie with the largest number of votes and work down. Aim to cover all movies with 10,000 or more motes.
  + Each Day
    - Update total vote counts.
    - Check for movies where the number of votes since the last scrape has increased by 1,000 or more. Scrape these first.
    - If scrape did not finish the previous day, continue the scrape.
* For all sets of movie votes,
  + Calculate the average vote
  + Fit the mixed model, and calculate
    - Polarity
    - Beta mean
    - Polar mean
* Perform analysis of movie polarity.