# Extracting the rest of the data

I applied the same process for extracting all PSP RPG data on the DS RPG's, SNES RPG's, PSX RPG's and PS2 RPG's.

# Cleaning the data

After loading my scraped data, I eliminated all the games that had at least 1 NA found in their game data. For fun and possibly for future reference, I created two new numerical variables:

Rating rate - the rate at which a GameFAQs user would rate the game

Difficulty(numeric) - just converting the difficulty ratings to numerical values on a scale of 5. 2/5 would be "Easy", 2.5/5 would be "Easy-Just Right" and so on.

# More exploratory data analysis

I made histograms of the rating, game length, and completion rate for each of the other systems.

# Trying K-means

I applied K-means clustering using rating, game length, and completion rate. I chose to have 4 clusters since there are 4 difficulty ratings, but the number of clusters can vary. I then made a 3d scatter plot and looked at which games corresponded to the purple dots.