D1: Central tendency

This week's assignment (D1) is very straightforward.

We would like you to write code that

1) Loads the dataset in 'movieDataReplicationSet.csv'. This dataset contains the movie rating data of 400 movies from 1097 research participants on a scale from 0 to 4. Missing values in this dataset represent movies that the participant declined to rate, for whatever reason (usually because they had not seen it).

2) Computes the mean rating of each movie, across all participants (the output here should be a variable that contains 400 means)

3) Computes the median rating of each movie, across all participants (the output here should be a variable that contains 400 medians)

4) Computes the modal rating of each movie, across all participants (the output here should be a variable that contains 400 modes)

5) Computes the mean of all the mean ratings from 2) to calculate a "mean of means"

Hint: Whereas you could type all of this out by hand, your life will be much easier if you use loops and/or vector operations. You can use the pre-existing mean, median and mode functions in Python. You don't have to write your own (but you can, if you want to).

Note: Because there is missing data, you'll need to find a way to handle those. There is a code session (CS) on that, which you could check out. It will also be covered in the lab.