- 1. You will get a movie review data set in this folder with two columns movie, review
- 2. Write a program that identifies top 3 pairs of movies that are very much alike
- 3. Hint:
  - a. Calculate movie review vector
  - b. Find similarity using
    - i. dot product formula followed by
    - ii. angle between two vector formula

Libraries to use: Only these two - Pandas for data read, numpy for vectors.

Write the dot product using numpy.dot() and compare with hand coded multiplication and addition as alternative. How much speedup do you see with numpy.dot()? Why?