```
% Tony Hyun Kim
% CS 246, PS 2, Problem 2(c)
function driver

clear all;
source = 'Matrix.txt';
A0 = load(source);

[U S V] = svd(A0,'econ');

for k = [197 247 297 347 397 447 497]
% for k = fliplr(297:10:497)
    % Apply compression
    mask = (1:497)<=k; mask = mask';
    Sk = diag(mask.*diag(S));
    A = U*Sk*V';

    % Compute the cosine similarity for the baseball docs
    sum_baseball = 0;
    for i = 1:99
        for j = (i+1):100
            sum_baseball = sum_baseball + cosim(A(i,:)',A(j,:)');
        end
    end
    avg_baseball = sum_baseball/nchoosek(100,2);

    % Compute the similarity among all documents
    sum_all = 0;
    for i = 1:496
        for j = (i+1):497
            sum_all = sum_all + cosim(A(i,:)',A(j,:)');
        end
    end
    avg_all = sum_all/nchoosek(497,2);

    % The 'r' score
    r = avg_baseball/avg_all;
    fprintf('%3d %.4f\n',k,r); % Dump output
end

% Expects u and v to be column vectors
function s = cosim(u,v)

s = (u'*v)/...
    (sqrt(u'*u)*sqrt(v'*v));
```