Question 1:

For question 1, we were to find 3 users that shared our age, gender and occupation. Those 3 users also had to print out their top 3 favorite and least favorite films. The code I used for this part was found on page 26 in the PCI textbook. I used users:

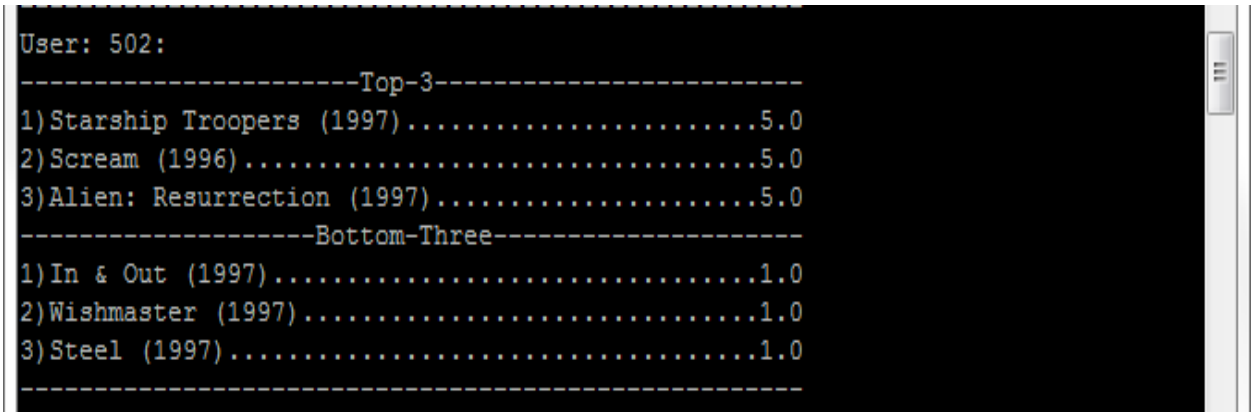
442
459
502

```
User: 442:
----------------------Top-3-------------------------
1)Star Trek: The Wrath of Khan (1982).............5.0
2)Better Off Dead... (1985).......................5.0
3)Escape from New York (1981).....................5.0
--------------------Bottom-Three--------------------
1)Postino, Il (1994)..............................1.0
2)Turbulence (1997)...............................1.0
3)Beautician and the Beast, The (1997).............1.0
-------------------------------------------------
```

User: 442:
----------Top-3----------
1)Star Trek: The Wrath of Khan (1982).....5.0
2)Better Off Dead... (1985).....5.0
3)Escape from New York (1981).....5.0
----------Bottom-Three----------
1)Postino, Il (1994).....1.0
2)Turbulence (1997).....1.0
3)Beautician and the Beast, The (1997).....1.0

```
User: 459:
----------------------Top-3-------------------------
1)Men in Black (1997).............................5.0
2)Cats Don't Dance (1997).........................5.0
3)Liar Liar (1997)................................5.0
--------------------Bottom-Three--------------------
1)Adventures of Pinocchio, The (1996).............1.0
2)Mars Attacks! (1996)............................1.0
3)Matilda (1996)..................................1.0
-------------------------------------------------
-------------------------------------------------
```

User: 459:
----------Top-3----------
1)Men in Black (1997).....5.0
2)Cats Don't Dance (1997).....5.0
3)Liar Liar (1997).....5.0
----------Bottom-Three----------
1)Adventures of Pinocchio, The (1996).....1.0
2)Mars Attacks! (1996).....1.0
3)Matilda (1996).....1.0

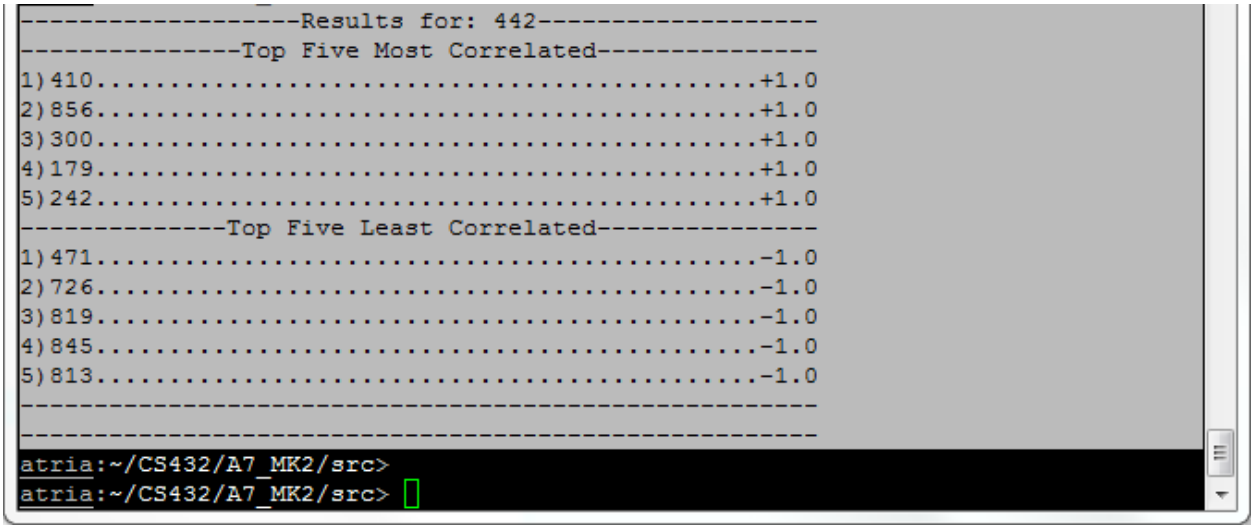


User: 502:
----------Top-3----------
1)Starship Troopers (1997).....5.0
2)Scream (1996).....5.0
3)Alien: Resurrection (1997).....5.0
----------Bottom-Three----------
1)In & Out (1997).....1.0
2)Wishmaster (1997).....1.0
3)Steel (1997).....1.0

Question 2:

For question 2, we were to choose 1 of the 3 users from question 1 and label them as a substitute of us. I chose user 442 as my substitute and found out which other users were mostly correlated and least correlated with my substitute.

442



Question 3:

For question 3, we were to compute the ratings for the films my subsitute hasn't seen as well give top and bottm 5 recommendations for films my subsitute should see. The function I used for this piece was on page 25 of the PCI textbook.

```
atria:~/CS432/A7_MK3> python part4.py
/usr/lib/python2.7/dist-packages/scipy/stats/stats.py:2514: RuntimeWarning: inva
lid value encountered in double_scalars
  r = r_num / r_den
Top 5 recommendations for films
Entertaining Angels: The Dorothy Day Story (1996)      Most likely rating:  5.0
Great Day in Harlem, A (1994)      Most likely rating:  5.0
They Made Me a Criminal (1939)      Most likely rating:  5.0
Prefontaine (1997)      Most likely rating:  5.0
Saint of Fort Washington, The (1993)      Most likely rating:  5.0

Bottom 5 recommendations for films
Speed (1994)      Most likely rating:  0
Dragonheart (1996)      Most likely rating:  0
Star Kid (1997)      Most likely rating:  0
Marlene Dietrich: Shadow and Light (1996)       Most likely rating:  0
James and the Giant Peach (1996)      Most likely rating:  0
atria:~/CS432/A7_MK3>
```

Question 4:

For question 4 we were to calculate the films that best and least correlated with the selection that I would personally make when it comes to movies. I agree with the results for I enjoy a majority of the films that best correlated with me, however I am neutral for the films that were correlated with the least best movies, for I do not recognize some of the titles.

The code function for this portion was taken from the recommended reading from PCI textbook.

```
pickedId='459'
pickedUser=linktable[pickedId]
for uid in linktable :
        if uid==pickedId:
                continue
        pickedUserRating=[]
        currentUserRating=[]
        for mid in pickedUser:
                if linktable[uid].has_key(mid) :
                        pickedUserRating.append(pickedUser[mid])
                        currentUserRating.append(linktable[uid][mid])
        if len(currentUserRating)==0 :
                correlation[uid]=0
        else:
                correlation[uid]=scipy.stats.pearsonr(pickedUserRating,currentUserRating)[0]
                if not correlation[uid] or math.isnan(correlation[uid]) :
```

```
        correlation[uid]=float(1)/(float(1)+scipy.spatial.distance.euclidean(pickedUserRating,currentUse
rRating))
        #calculate estimated rating
        for mid in linktable[uid]:
                # ignore scores of zero or lower and only score movies I haven't seen yet
                if mid not in pickedUser and  correlation[uid]>0:
                        movies[mid]['wtotal']+=linktable[uid][mid]*correlation[uid]
                        movies[mid]['stoal']+=correlation[uid]
#calculate rating
for m in movies:
        if  movies[m]['stoal']!= 0:
                movies[m]['erate']=float(movies[m]['wtotal'])/float(movies[m]['stoal'])
        # if movies[m]['erate']>5 or movies[m]['erate']< -5:
        #         print movies[m]
```

```
atria:~/CS432/A7_MK3> python part34.py
favorate move: 257|Men in Black (1997)|04-Jul-1997|
/usr/lib/python2.7/dist-packages/scipy/stats/stats.py:2514: RuntimeWarning: inva
lid value encountered in double_scalars
  r = r_num / r_den
Top 5 most correlated movies:
Wife, The (1995)  ( correlation: 1.0 )
Nico Icon (1995)  ( correlation: 1.0 )
Letter From Death Row, A (1998)  ( correlation: 1.0 )
Before the Rain (Pred dozhdot) (1994)  ( correlation: 1.0 )
Savage Nights (Nuits fauves, Les) (1992)  ( correlation: 1.0 )
Bottom 5 least correlated movies:
Month by the Lake, A (1995) ( correlation: -1.0 )
Joy Luck Club, The (1993) ( correlation: -1.0 )
Bewegte Mann, Der (1994) ( correlation: -1.0 )
8 Seconds (1994) ( correlation: -1.0 )
Endless Summer 2, The (1994) ( correlation: -1.0 )

Least favorate move: 254|Batman & Robin (1997)|20-Jun-1997|
Top 5 most correlated movies:
Apostle, The (1997)  ( correlation: 1.0 )
Once Upon a Time... When We Were Colored (1995)  ( correlation: 1.0 )
Farinelli: il castrato (1994)  ( correlation: 1.0 )
Once Were Warriors (1994)  ( correlation: 1.0 )
Welcome To Sarajevo (1997)  ( correlation: 1.0 )
Bottom 5 least correlated movies:
Contempt (Mépris, Le) (1963) ( correlation: -1.0 )
Mixed Nuts (1994) ( correlation: -1.0 )
Fluke (1995) ( correlation: -1.0 )
Portrait of a Lady, The (1996) ( correlation: -1.0 )
Search for One-eye Jimmy, The (1996) ( correlation: -1.0 )
atria:~/CS432/A7_MK3> python part4.py
```

Top 5 recommendations for films
Entertaining Angels: The Dorothy Day Story (1996) Most likely rating: 5.0
Great Day in Harlem, A (1994) Most likely rating: 5.0
They Made Me a Criminal (1939) Most likely rating: 5.0
Prefontaine (1997) Most likely rating: 5.0
Saint of Fort Washington, The (1993) Most likely rating: 5.0

Bottom 5 recommendations for films
Speed (1994) Most likely rating: 0
Dragonheart (1996) Most likely rating: 0
Star Kid (1997) Most likely rating: 0
Marlene Dietrich: Shadow and Light (1996) Most likely rating: 0
James and the Giant Peach (1996) Most likely rating: 0