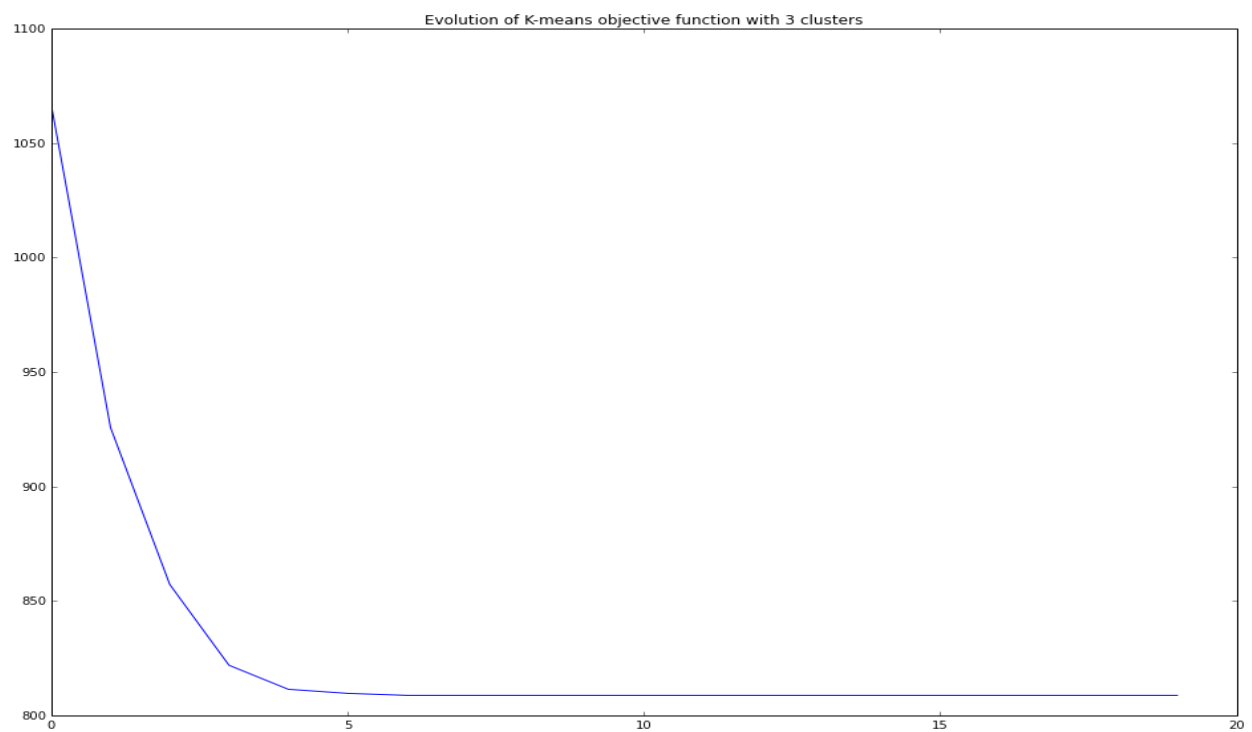
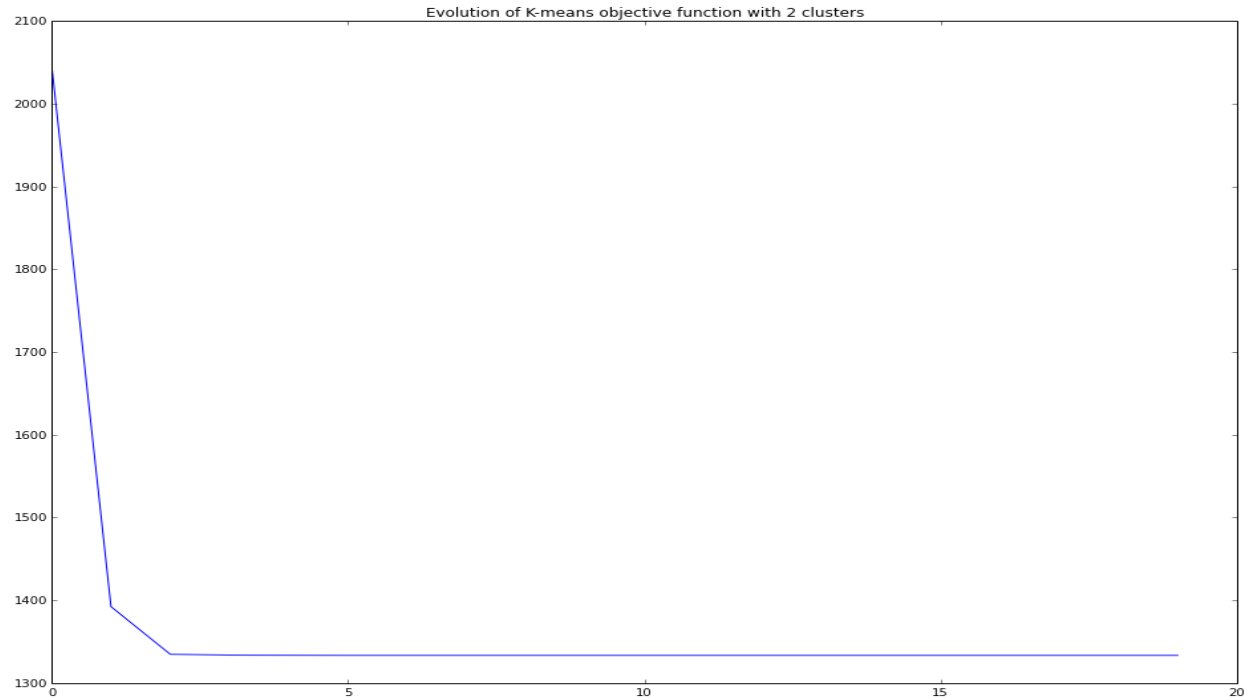


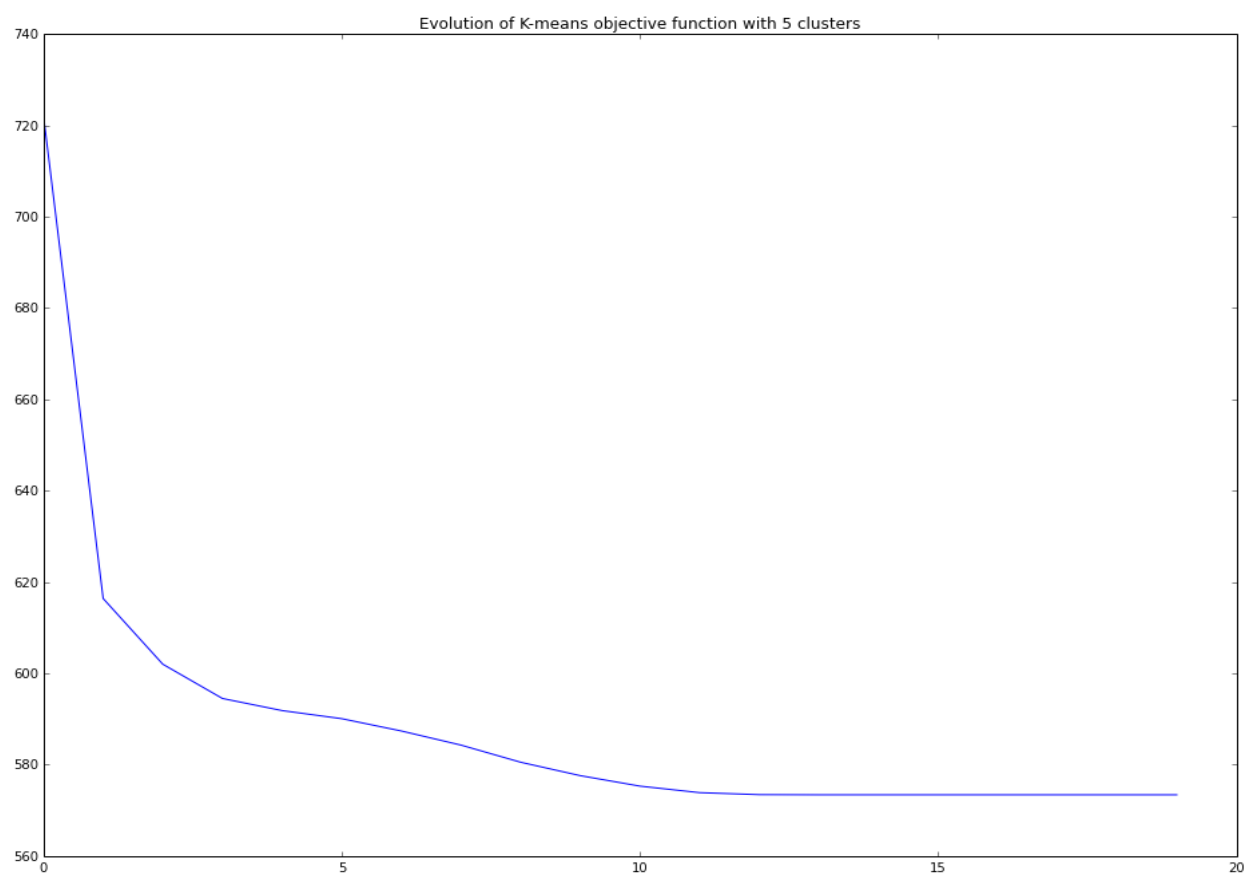
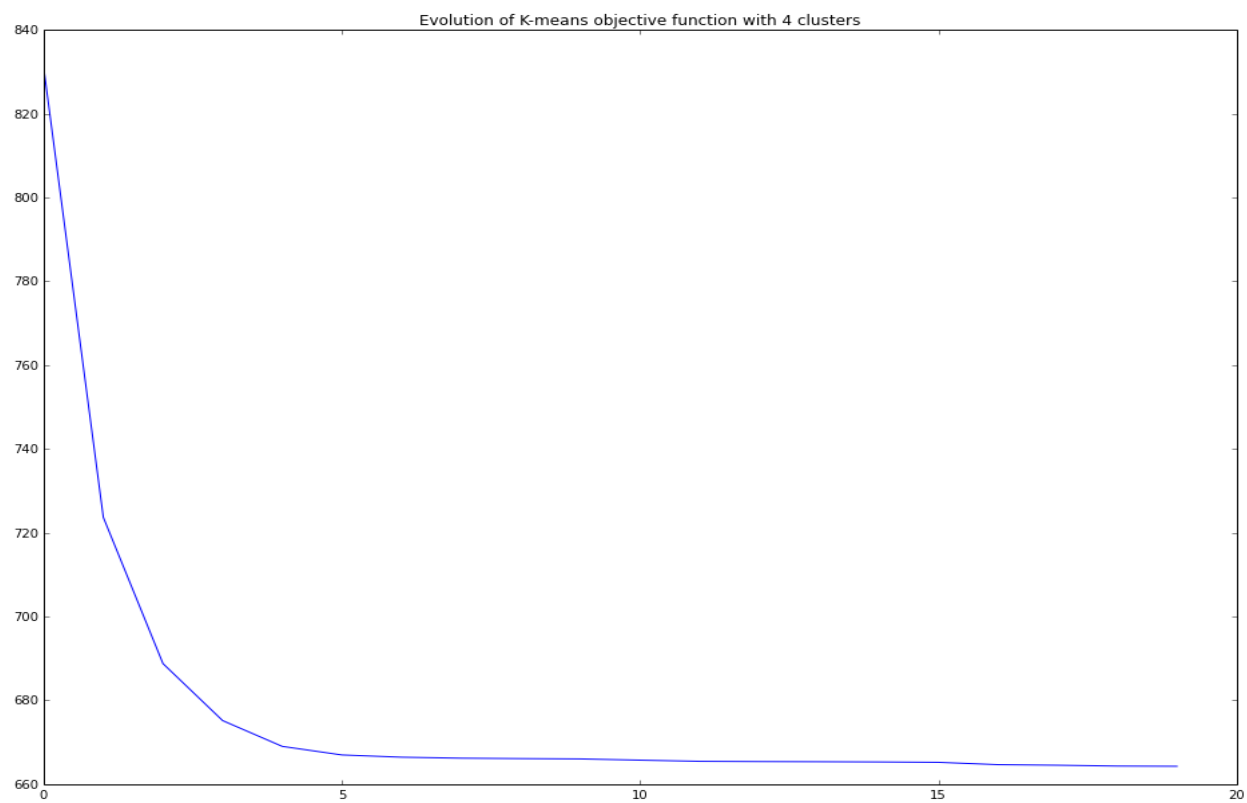
Machine Learning (HW4)

Sirui Tan, st2957

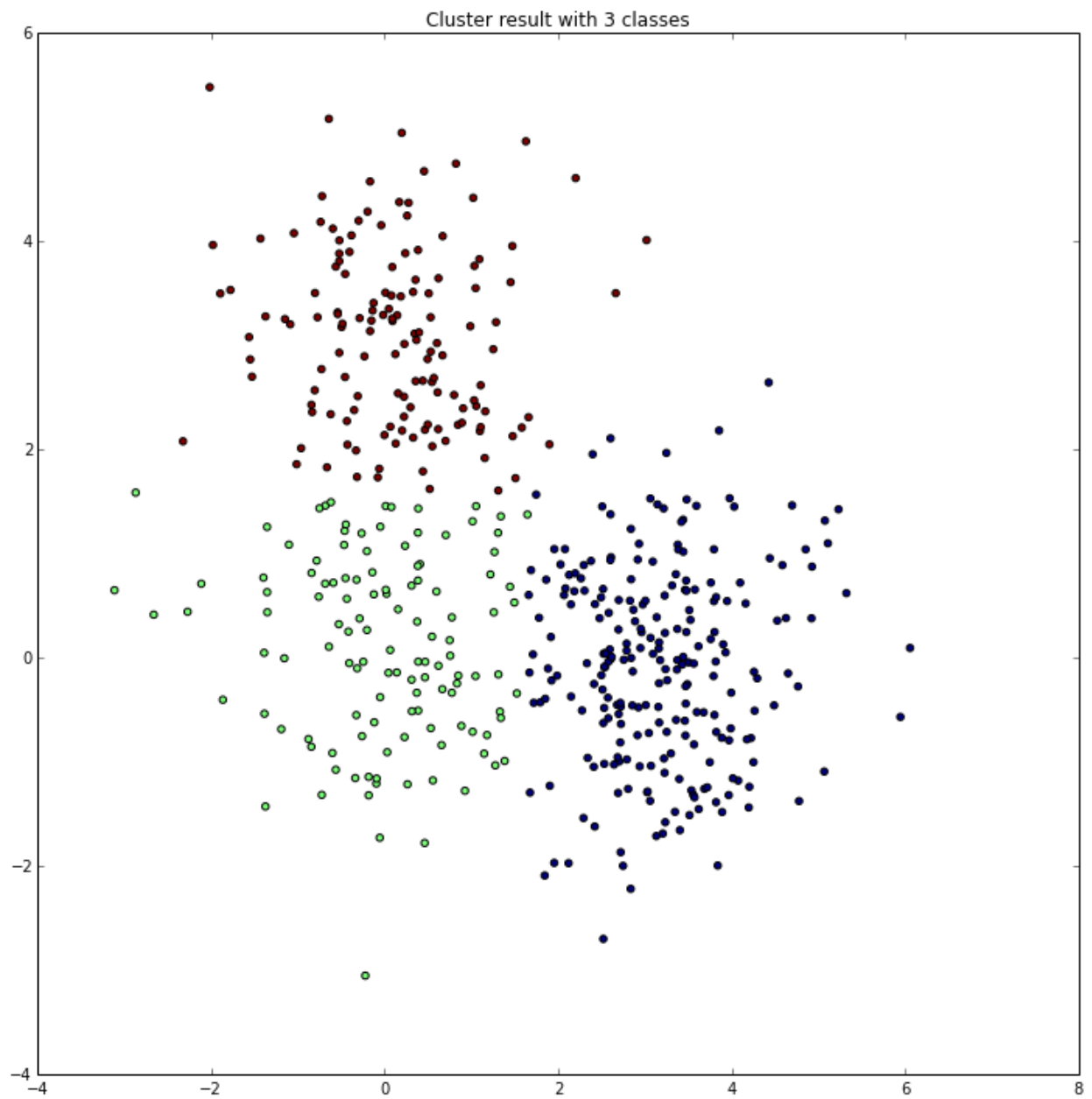
Problem 1

Objective functions:





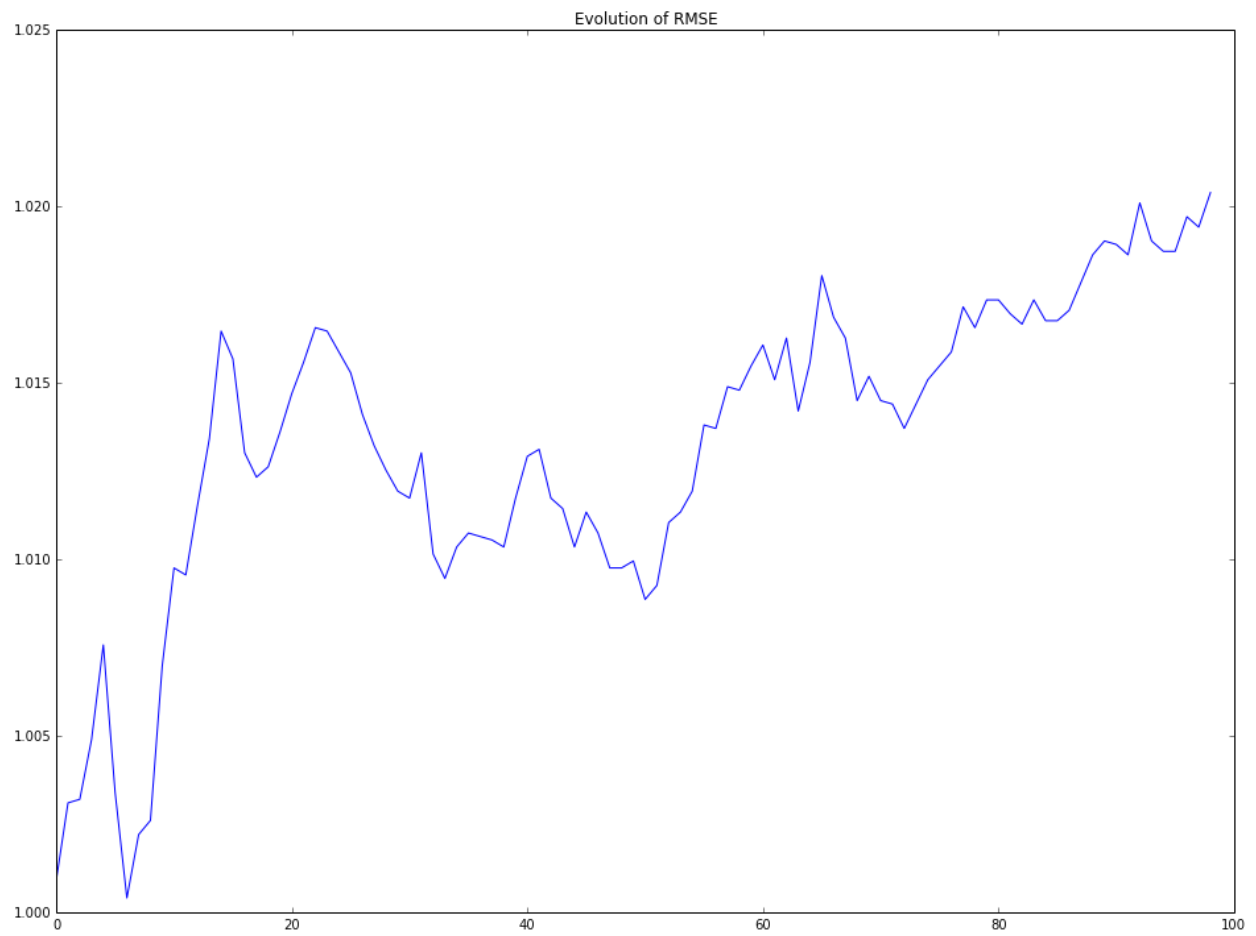
Cluster results:



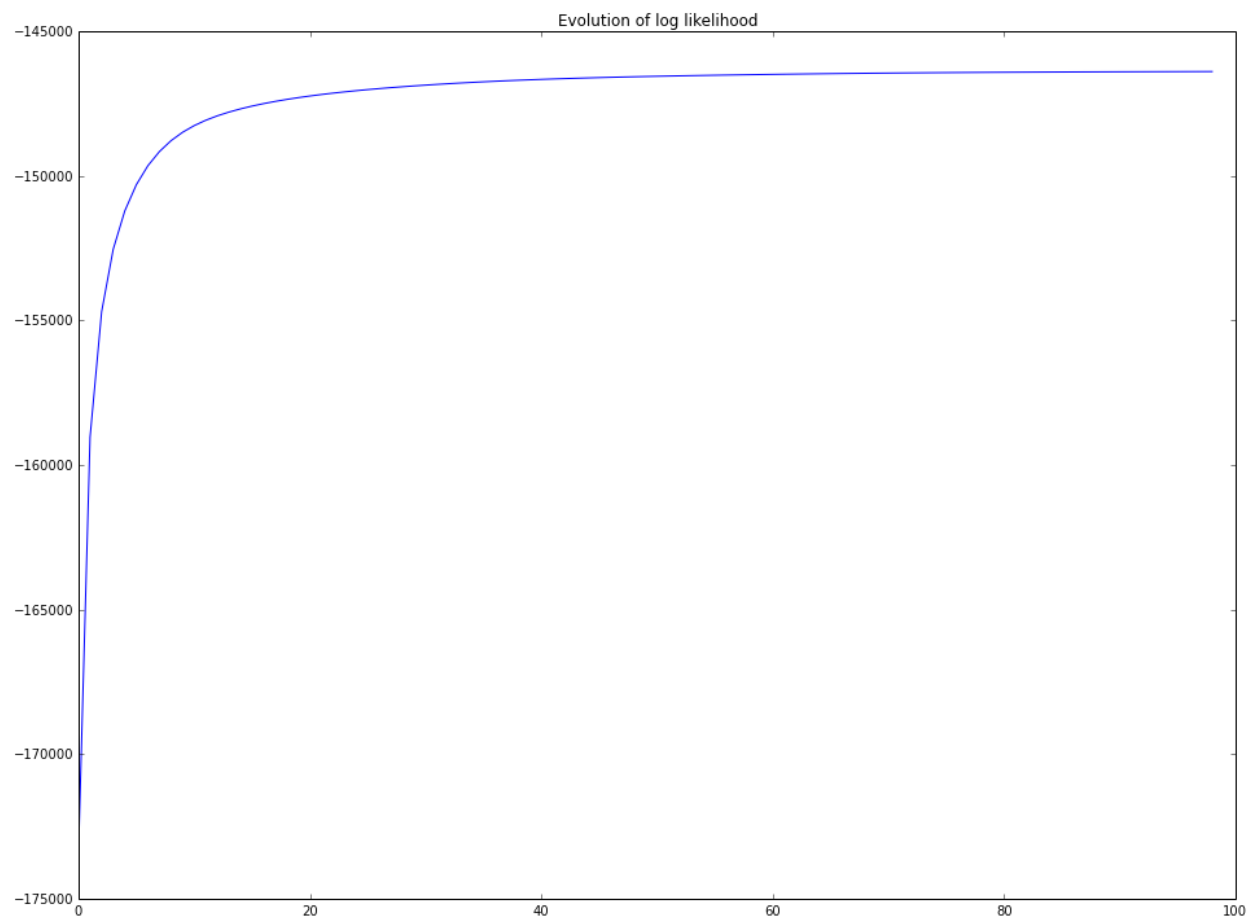


Problem 2

RMSE:



Log Likelihood:



Movie vicinities:

Toy Story (1995)	Babe (1995)	0.619304121769
	Return of the Pink Panther, The (1974)	0.669550568157
	Indiana Jones and the Last Crusade (1989)	0.681705397292
	Aladdin (1992)	0.683822132479
	Searching for Bobby Fischer (1993)	0.711210658208
Star Wars (1977)	Empire Strikes Back, The (1980)	0.276554634511
	Return of the Jedi (1983)	0.501893965932
	Raiders of the Lost Ark (1981)	0.686696193527
	Blues Brothers, The (1980)	0.791751757796
	Star Trek: The Wrath of Khan (1982)	0.898083425627
Roman Holiday (1953)	North by Northwest (1959)	0.529803774235
	Casablanca (1942)	0.532199356626
	To Catch a Thief (1955)	0.537460384472
	Snow White and the Seven Dwarfs (1937)	0.561054565061
	Manchurian Candidate, The (1962)	0.565022347391

Centroids with the most users:

1

- **# of users:** 88

Coordinate:

[-0.81220154 0.57874215 0.27862718 0.05765165 0.70975647 -1.07124366
-0.77626583 0.51513332 0.08898882 -0.63622615]

Top 10:

Silence of the Lambs, The (1991) 4.45727367194

Godfather, The (1972) 4.60195372916

12 Angry Men (1957) 4.49009468134

Schindler's List (1993) 4.53617209331

One Flew Over the Cuckoo's Nest (1975) 4.55787745824

Close Shave, A (1995) 4.49092220829

Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb (1963) 4.46617203014
Casablanca (1942) 4.57125416854
Lawrence of Arabia (1962) 4.64677408089
Rear Window (1954) 4.50227788595

2

- **# of users:** 73.0
- **Coordinates:**
[-0.71130771 0.01376001 0.33042673 0.05157048 1.05185977 -1.05403351
-0.39714222 0.68007137 0.54132103 -0.27907765]
- **Top 10:**
Usual Suspects, The (1995) 4.51172102521
Braveheart (1995) 4.31293137413
Shawshank Redemption, The (1994) 4.42308786535
Much Ado About Nothing (1993) 4.32080957212
Princess Bride, The (1987) 4.28720230859
Henry V (1989) 4.48337827456
Good Will Hunting (1997) 4.32983576948
L.A. Confidential (1997) 4.32059575332
Schindler's List (1993) 4.38497628843
Boot, Das (1981) 4.32459296187

3

- **# of users:** 64.0
- **Coordinates:**
[-0.75006375 0.602916 0.457247 0.28307571 1.21303291 -0.39644043
-0.36198851 0.81328807 0.35378007 -0.46344017]
- **Top 10:**
Usual Suspects, The (1995) 4.57122054604
Hoop Dreams (1994) 4.33150846182
Pulp Fiction (1994) 4.57645207195
Wallace & Gromit: The Best of Aardman Animation (1996) 4.52691390904
Wrong Trousers, The (1993) 4.48123500044
Chasing Amy (1997) 4.31745039659
Good Will Hunting (1997) 4.32833133304
L.A. Confidential (1997) 4.50327926876
Close Shave, A (1995) 4.33989245056
Ghost in the Shell (Kokaku kidotai) (1995) 4.34058118131

4

- **# of users:** 63.0
- **Coordinates:**
[-0.52035147 0.35378268 0.92595641 0.14892151 0.89567213 -0.79366501
-1.02880906 0.70122522 0.12811957 -0.5294529]

- **Top 10:**
 Three Colors: Red (1994) 4.4044972693
 Three Colors: Blue (1993) 4.49447416566
 Fargo (1996) 4.48872680591
 Wallace & Gromit: The Best of Aardman Animation (1996) 4.4860721051
 Lone Star (1996) 4.439172178
 Big Night (1996) 4.41668352886
 Clockwork Orange, A (1971) 4.62056525855
 Kolya (1996) 4.46768389595
 Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb (1963) 4.70263257164
 Wings of Desire (1987) 4.50170705971

5

- **# of users:** 62.0
- **Coordinates:**
 [-0.31323655 0.41981868 0.58567958 0.87976441 0.96395996 -1.06203204
 -0.54490062 0.91917021 0.45774257 -0.59606361]
- **Top 10:**
 Star Wars (1977) 4.45514804449
 Shawshank Redemption, The (1994) 4.45183872006
 Godfather, The (1972) 4.46722884184
 Citizen Kane (1941) 4.43195417767
 Chasing Amy (1997) 4.56552491502
 In the Company of Men (1997) 4.50408016345
 Wag the Dog (1997) 4.56301657939
 Casablanca (1942) 4.47775760765
 Kundun (1997) 4.43928222712
 Some Folks Call It a Sling Blade (1993) 4.54300466615

Centroids with the most movies:

1

- **# of movies:** 159.0
- **Coordinates:**
 [-0.02651252 0.28932827 0.11422134 0.19107266 0.3618944 -0.53029474
 -0.30826029 0.36464409 0.23806045 -0.19131528]
- **Top 10:**
 Substance of Fire, The (1996) 0.277298537561
 Original Gangstas (1996) 0.337718251969
 Guilty as Sin (1993) 0.30449099132
 Lover's Knot (1996) 0.260383296524
 Turning, The (1992) 0.330863884262
 I Like It Like That (1994) 0.30068949265
 Outlaw, The (1943) 0.330450478915

Jerky Boys, The (1994) 0.319601454199
Angela (1995) 0.263246169371
All Things Fair (1996) 0.247629315802

2

- **# of movies:** 140.0
- **Coordinates:**
[-0.30328056 0.38919958 0.38991863 0.14332636 0.49677395 -0.68484946
-0.46932198 0.56540928 -0.07191711 -0.32885448]
- **Top 10:**
Wild Bill (1995) 0.408160482565
Man of the Year (1995) 0.359943477753
Faces (1968) 0.382612346401
Last Time I Saw Paris, The (1954) 0.403705982628
Penny Serenade (1941) 0.376693713689
Angel and the Badman (1947) 0.291055719661
Nelly & Monsieur Arnaud (1995) 0.383638536569
Secret Adventures of Tom Thumb, The (1993) 0.404209150843
Spirits of the Dead (Tre passi nel delirio) (1968) 0.414055148578
Price Above Rubies, A (1998) 0.378051889872

3

- **# of movies:** 134.0
- **Coordinates:**
[-0.09800215 0.08407704 0.00107372 0.12586565 0.10493946 -0.15922827
-0.01976782 0.0984694 0.12509309 -0.0821539]
- **Top 10:**
Police Story 4: Project S (Chao ji ji hua) (1993) 0.177203588333
Homage (1995) 0.185873342021
Low Life, The (1994) 0.185873342021
Mille bolle blu (1993) 0.185873342021
Shadow of Angels (Schatten der Engel) (1976) 0.185873342021
Venice/Venice (1992) 0.17491027688
Bird of Prey (1996) 0.185873342021
Johnny 100 Pesos (1993) 0.185873342021
JLG/JLG - autoportrait de d♠cembre (1994) 0.185873342021
Temptress Moon (Feng Yue) (1996) 0.175406460817

4

- **# of movies:** 107.0
- **Coordinates:**
[-0.08670361 0.1401618 -0.24240753 0.4787516 0.57375253 -0.70889747
-0.08438435 0.77248156 0.17200016 -0.21513649]
- **Top 10:**
Sgt. Bilko (1996) 0.482410732318
Batman & Robin (1997) 0.460500731993

Made in America (1993) 0.505495516666
I Love Trouble (1994) 0.469667264548
Twisted (1996) 0.454779986801
Specialist, The (1994) 0.503535211948
Moonlight and Valentino (1995) 0.458393249521
Scarlet Letter, The (1995) 0.423657514386
Visitors, The (Visiteurs, Les) (1993) 0.456840620035
Trial by Jury (1994) 0.513661699283

5

- **# of movies:** 105.0
- **Coordinates:**
[-0.68554561 0.3398665 0.04338228 -0.02954462 0.75271241 -1.13923248
-0.76571332 0.72288537 0.19274382 -0.55766827]
- **Top 10:**
20,000 Leagues Under the Sea (1954) 0.41702538693
12 Angry Men (1957) 0.452580072061
Casablanca (1942) 0.442631626793
Roman Holiday (1953) 0.489462602778
To Catch a Thief (1955) 0.32618852061
Thin Man, The (1934) 0.467910021166
African Queen, The (1951) 0.413250763978
American in Paris, An (1951) 0.479087863329
Rear Window (1954) 0.479784478262
Arsenic and Old Lace (1944) 0.413477930517