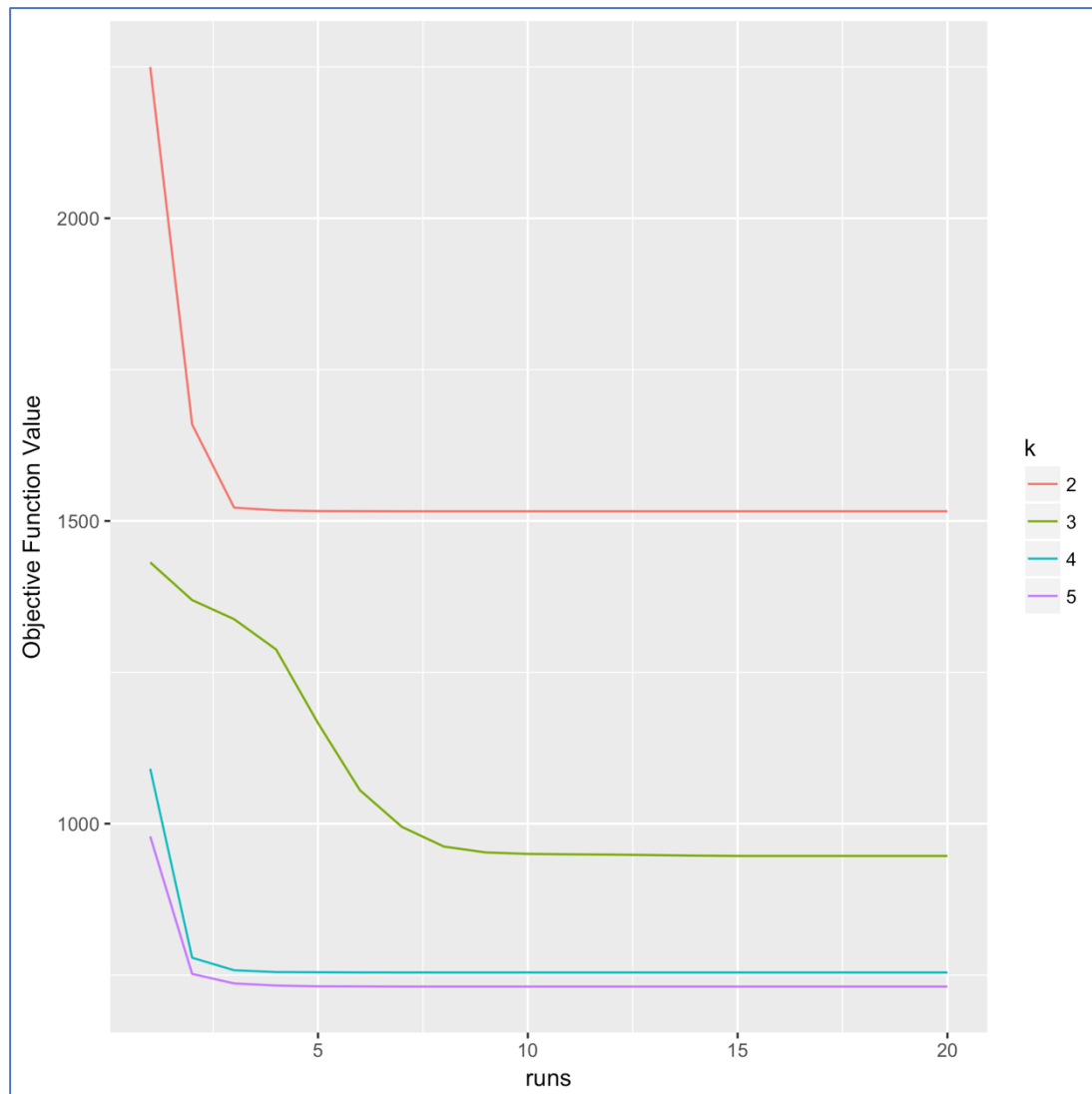


Question 1

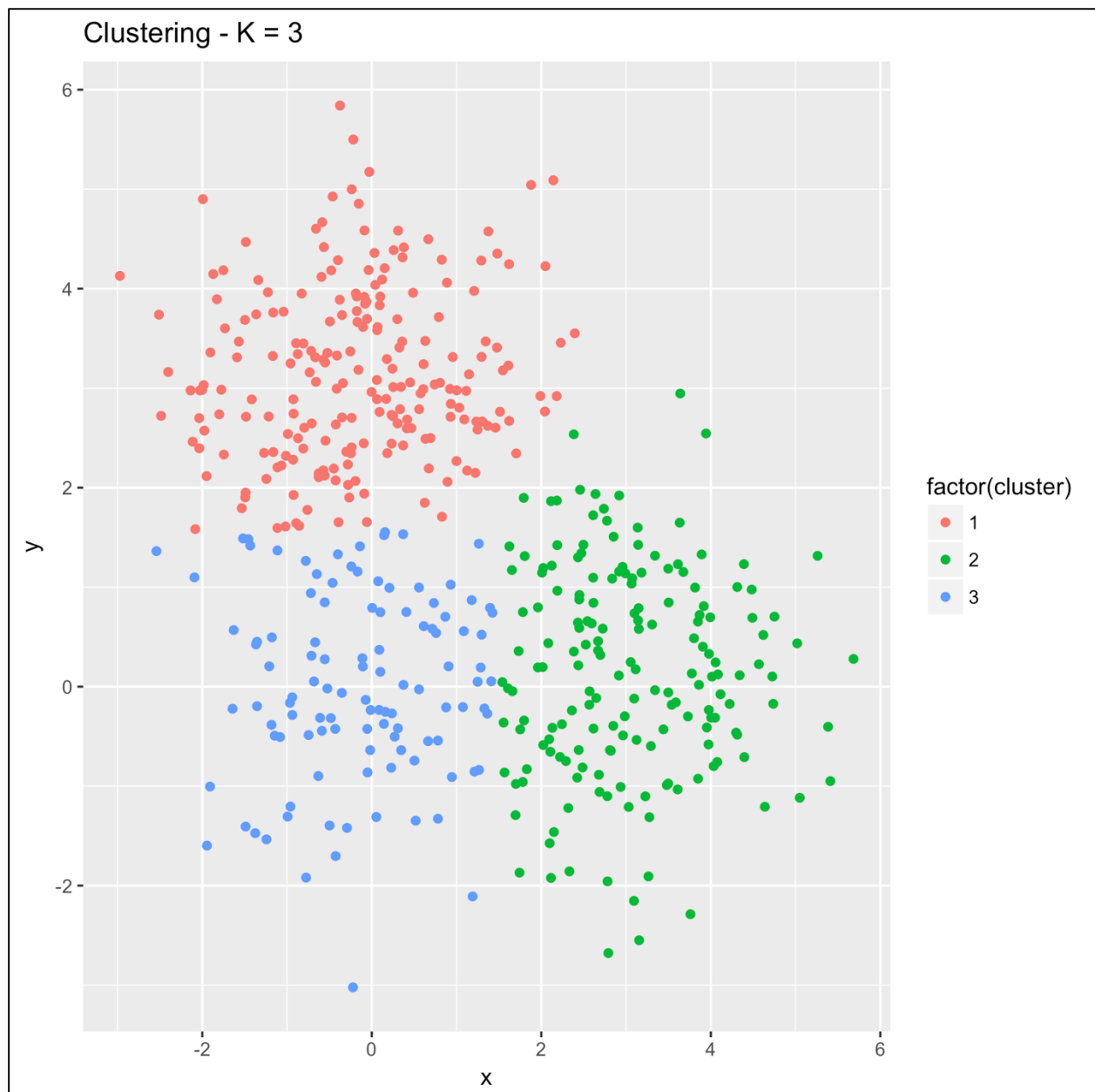
a)

For $K = 2, 3, 4, 5$, we plot the value of the K-means objective function per iteration for 20 iterations. We notice that the objective function converges to its optimal value much before the 20th iteration in most cases.

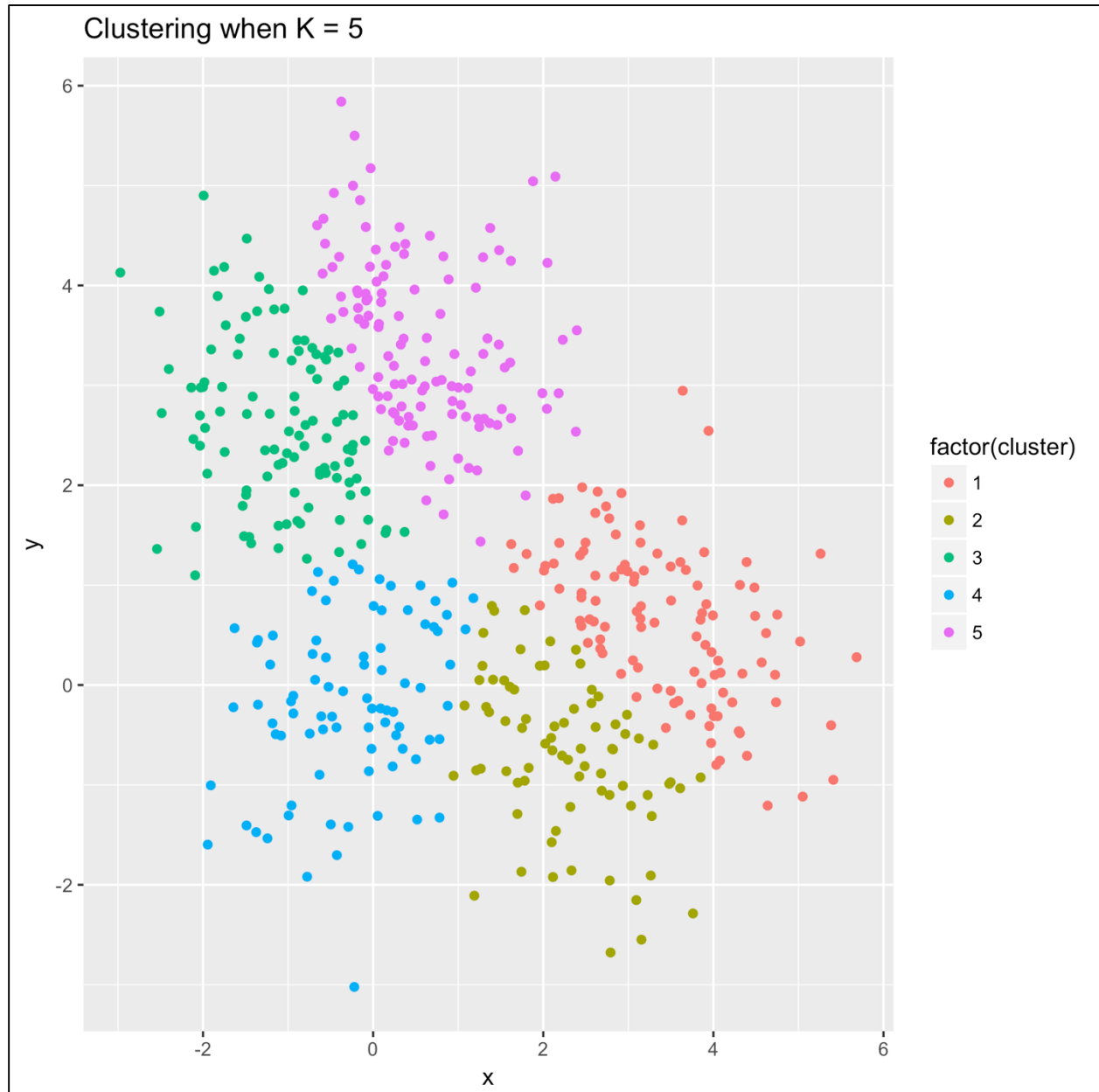


b)

For $K = 3$, we plot the 500 data points and indicate the cluster of each for the final iteration by marking it with a color.



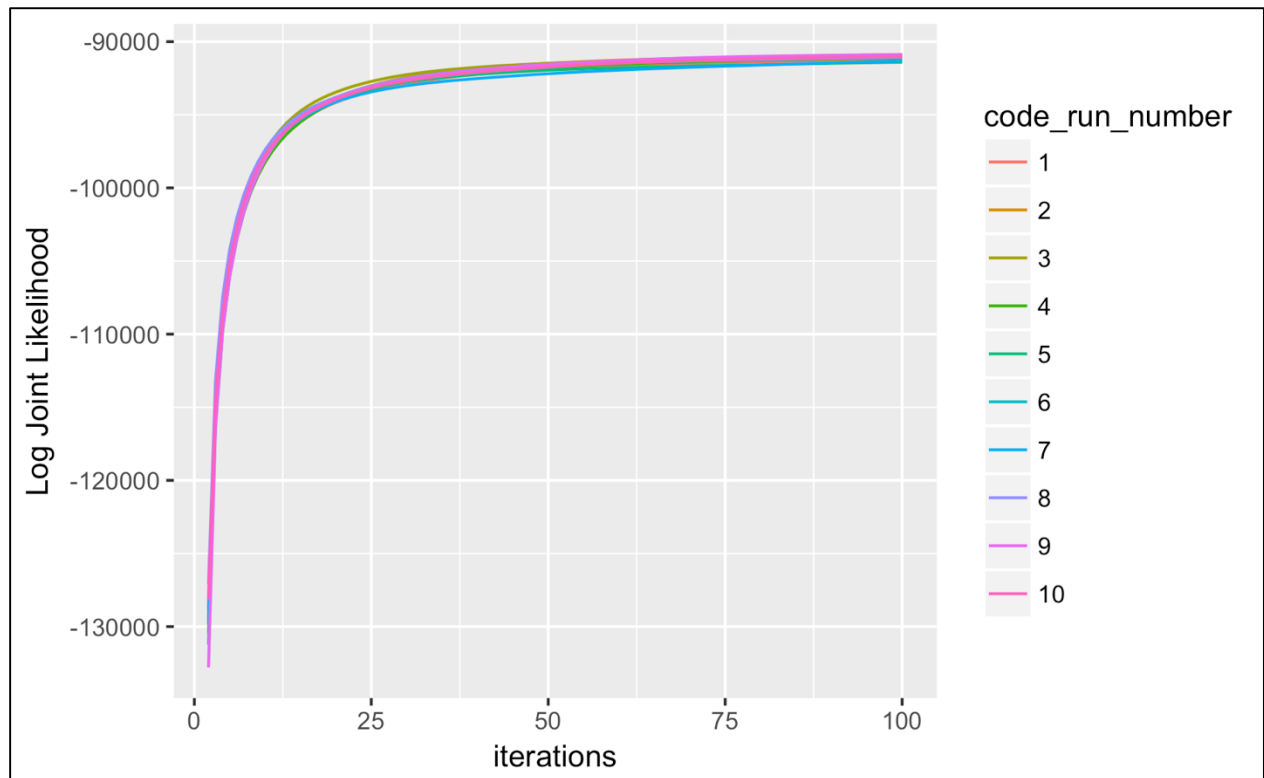
For $K = 5$, we plot the 500 data points and indicate the cluster of each for the final iteration by marking it with a color. There are two additional clusters on the graph now.



Question 2

a)

Notice, in the graph below, we plot the objective function for **iterations 2 to 100**. The objective value for the first iteration was much lower, because of the prior we have on all the U and V vectors, when we initialize them.



Code_run_number	final_obj_function_values (training)	RMSE (test set)
9	-90874.39868	1.114682242
2	-90887.27796	1.152843288
6	-90952.59631	1.106755088
10	-91014.45612	1.130057718
8	-91071.05162	1.101928957
3	-91092.43098	1.159538852
4	-91123.4782	1.10986428
1	-91235.57247	1.162652622
7	-91314.37974	1.169402581
5	-91406.70008	1.157038221

b)

Following is the list of our query movies with their 11 nearest neighbours (each movie is also its closest neighbor), ranked in ascending order of distances

Star Wars			
neighbour rank	movie_id	movie_name	distances
1	50	Star Wars (1977)	0
2	172	Empire Strikes Back, The (1980)	0.302859973
3	181	Return of the Jedi (1983)	0.607162077
4	174	Raiders of the Lost Ark (1981)	0.662458744
5	12	Usual Suspects, The (1995)	0.820627115
6	429	Day the Earth Stood Still, The (1951)	0.890037748
7	210	Indiana Jones and the Last Crusade (1989)	0.903695934
8	613	My Man Godfrey (1936)	0.933982938
9	186	Blues Brothers, The (1980)	0.948132204
10	195	Terminator, The (1984)	0.957826628
11	168	Monty Python and the Holy Grail (1974)	0.972849904

My Fair Lady			
neighbour rank	movie_id	movie_name	distances
1	485	My Fair Lady (1964)	0
2	486	Sabrina (1954)	0.674907872
3	465	Jungle Book, The (1994)	0.750726145
4	419	Mary Poppins (1964)	0.771721578
5	638	Return of Martin Guerre, The (Retour de Martin Guerre, Le) (1982)	0.796628835
6	8	Babe (1995)	0.845711871
7	602	American in Paris, An (1951)	0.857392999
8	143	Sound of Music, The (1965)	0.871493728
9	507	Streetcar Named Desire, A (1951)	0.893317201
10	945	Charade (1963)	0.916098579
11	132	Wizard of Oz, The (1939)	0.921864182

GoodFellas			
neighbour rank	movie_id	movie_name	distances
1	182	GoodFellas (1990)	0
2	646	Once Upon a Time in the West (1969)	0.483673337
3	693	Casino (1995)	0.592917188
4	504	Bonnie and Clyde (1967)	0.603306952
5	177	Good, The Bad and The Ugly, The (1966)	0.674741153
6	180	Apocalypse Now (1979)	0.731343129
7	508	People vs. Larry Flynt, The (1996)	0.762463907
8	187	Godfather: Part II, The (1974)	0.784780234
9	188	Full Metal Jacket (1987)	0.789481721
10	469	Short Cuts (1993)	0.862864721
11	135	2001: A Space Odyssey (1968)	0.867489956