Datasets:

Downloaded the binary image files for this homework from

The CIFAR-10 dataset consists of 60000 32x32 colour images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images.

The dataset is divided into five training batches and one test batch, each with 10000 images. The test batch contains exactly 1000 randomly-selected images from each class. The training batches contain the remaining images in random order, but some training batches may contain more images from one class than another. Between them, the training batches contain exactly 5000 images from each class.

Part-1 - Plot:
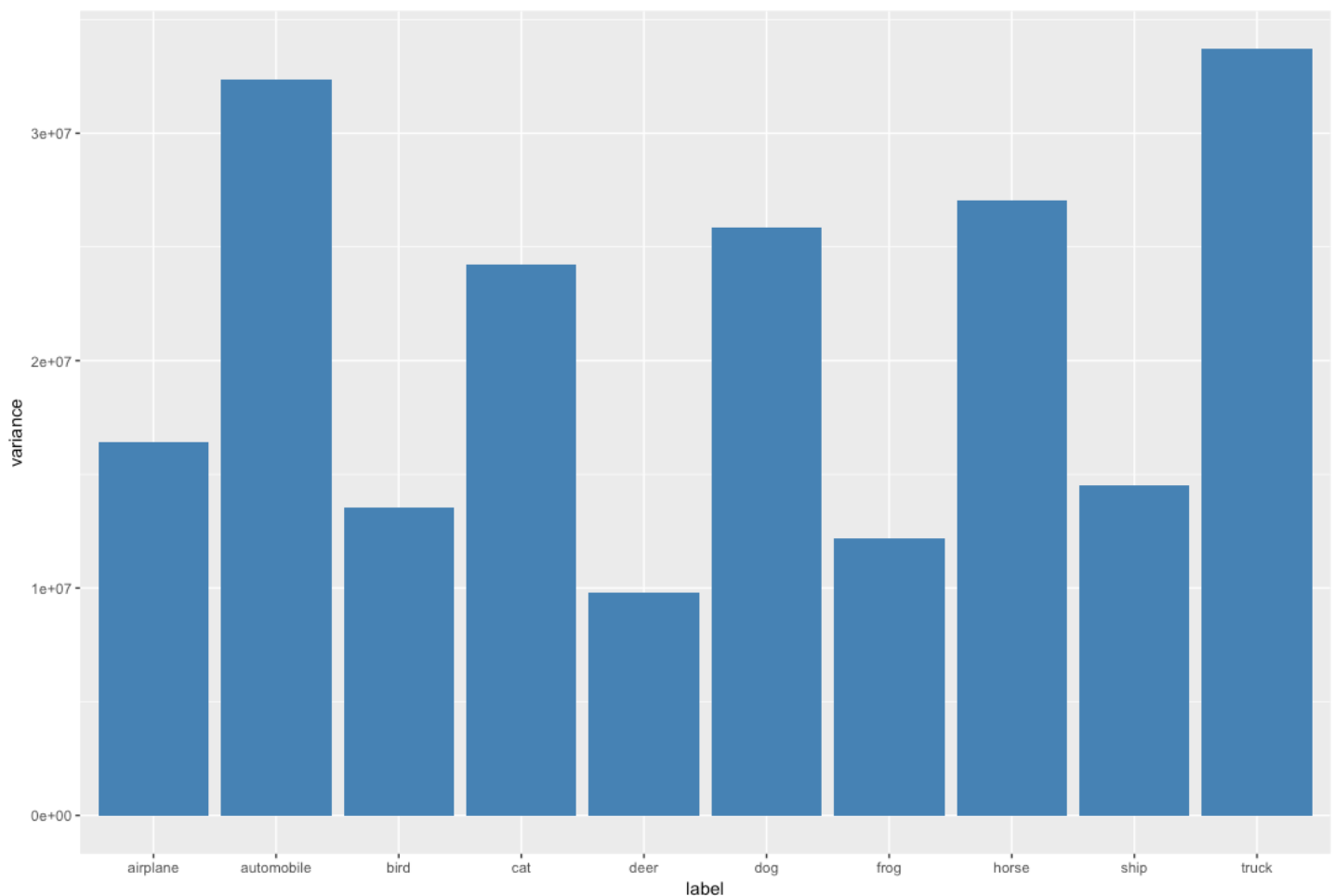
Part-2:

Distance Matrix:

Categories	airplane	automobile	bird	cat	deer	dog	frog	horse	ship	truck
airplane	0	1683.63535	1605.02435	1905.53526	2148.76341	1965.22149	2445.67973	1663.64593	945.541104	1449.09491
automobile	1683.63535	0	886.23675	1027.64978	1143.08137	1216.07943	1191.192	950.786078	1303.46655	949.995771
bird	1605.02435	886.23675	0	517.311502	601.250335	701.468226	913.747516	418.276306	1557.71502	1416.67473
cat	1905.53526	1027.64978	517.311502	0	469.791716	412.181669	677.491969	596.376737	1851.2145	1676.46786
deer	2148.76341	1143.08137	601.250335	469.791716	0	617.697141	460.510929	684.346911	2065.62166	1830.74085
dog	1965.22149	1216.07943	701.468226	412.181669	617.697141	0	828.581051	843.672091	1897.59182	1880.24377
frog	2445.67973	1191.192	913.747516	677.491969	460.510929	828.581051	0	948.704021	2249.19978	1913.24088
horse	1663.64593	950.786078	418.276306	596.376737	684.346911	843.672091	948.704021	0	1660.26808	1347.3341
ship	945.541104	1303.46655	1557.71502	1851.2145	2065.62166	1897.59182	2249.19978	1660.26808	0	1066.94163
truck	1449.09491	949.995771	1416.67473	1676.46786	1830.74085	1880.24377	1913.24088	1347.3341	1066.94163	0

Plot:
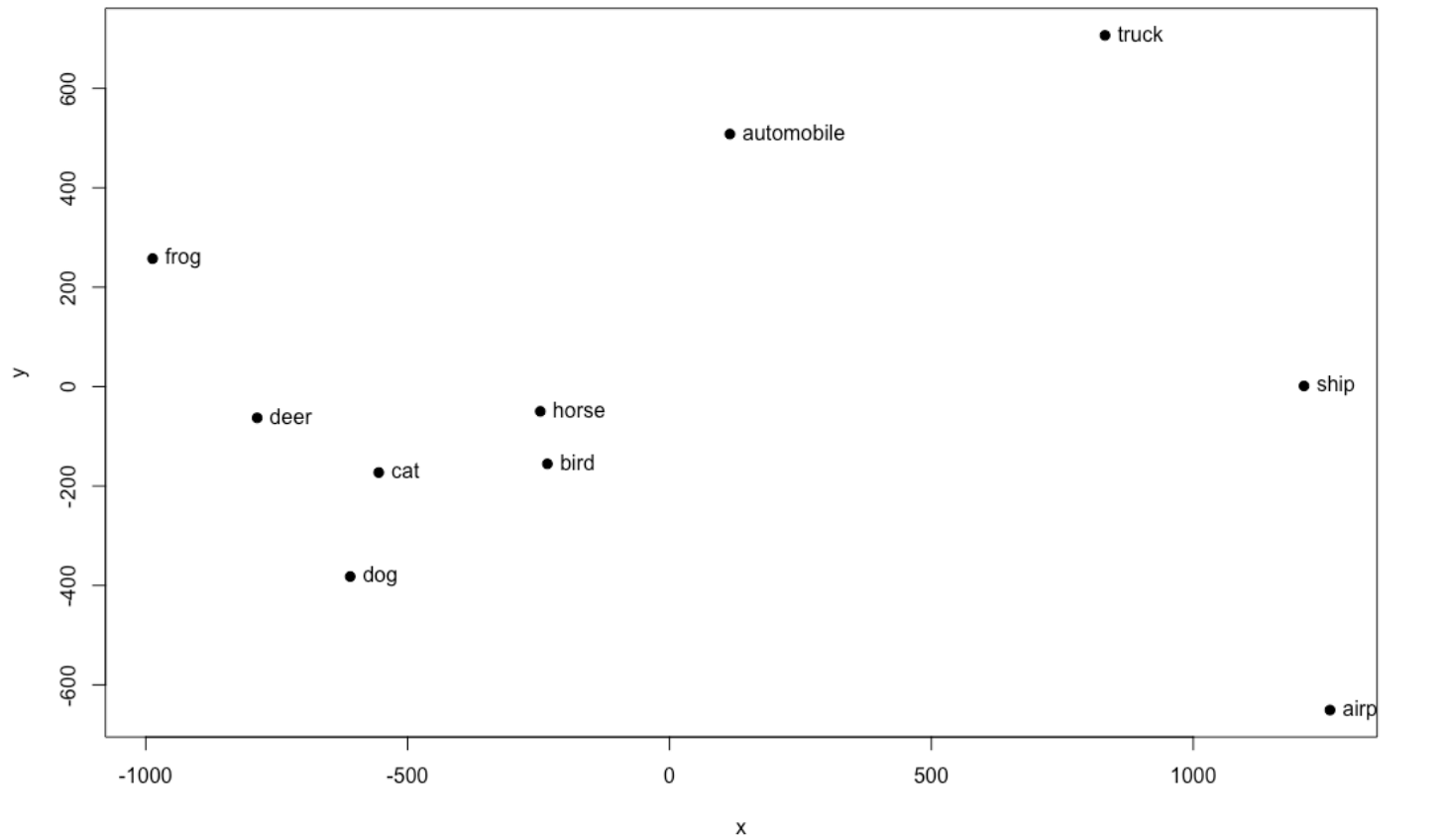
Network Graph
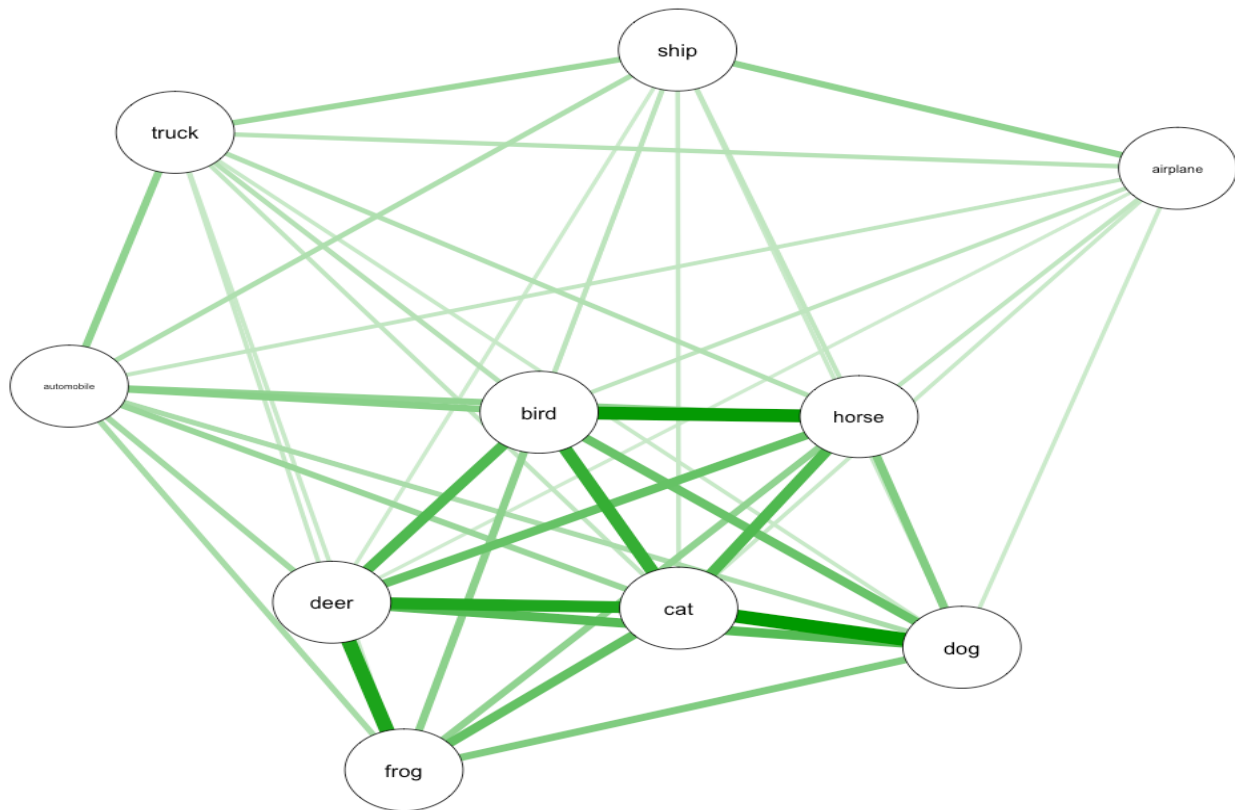
Part-3:

Distance Matrix:

Categories	airplane	automobile	bird	cat	deer	dog	frog	horse	ship	truck
airplane	3331788.15	4577527.76	3458508.26	4577134.58	3350845.72	4652084.83	3630453.16	4428601.78	3693497.01	4732702.98
automobile	4577527.76	4572350.31	4421248.52	5237038.06	4189306.64	5407481.61	4359446.81	5124795.63	4317245.85	4878578.43
bird	3458508.26	4421248.52	2916671.92	4040544.83	3050214.02	3983704.49	3224236.59	4093207.16	3667422.73	4532279.08
cat	4577134.58	5237038.06	4040544.83	4714462.21	4001883.41	4686489.62	4057071.85	4732205.5	4453617.73	5051417.03
deer	3350845.72	4189306.64	3050214.02	4001883.41	2824910.2	3987176.08	3189388.28	3853819.19	3353937.97	4231220.96
dog	4652084.83	5407481.61	3983704.49	4686489.62	3987176.08	4341686.96	4137389.44	4623033.27	4524100.49	5130355.72
frog	3630453.16	4359446.81	3224236.59	4057071.85	3189388.28	4137389.44	3115869.09	4181539.51	3691590.93	4427806.25
horse	4428601.78	5124795.63	4093207.16	4732205.5	3853819.19	4623033.27	4181539.51	4262915.6	4307932.14	4976368.13
ship	3693497.01	4317245.85	3667422.73	4453617.73	3353937.97	4524100.49	3691590.93	4307932.14	3302096.01	4345403.27
truck	4732702.98	4878578.43	4532279.08	5051417.03	4231220.96	5130355.72	4427806.25	4976368.13	4345403.27	4651158.94

Plot:
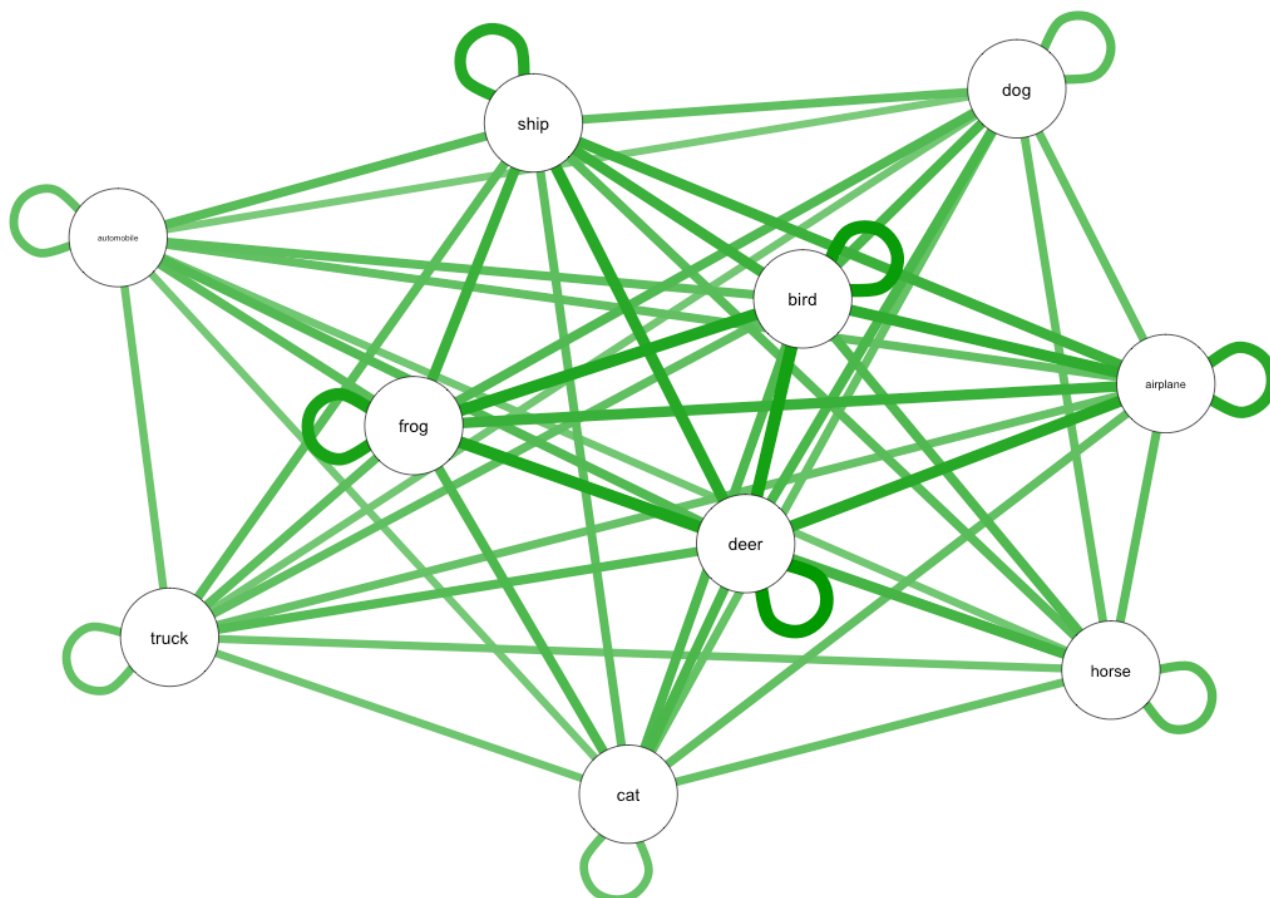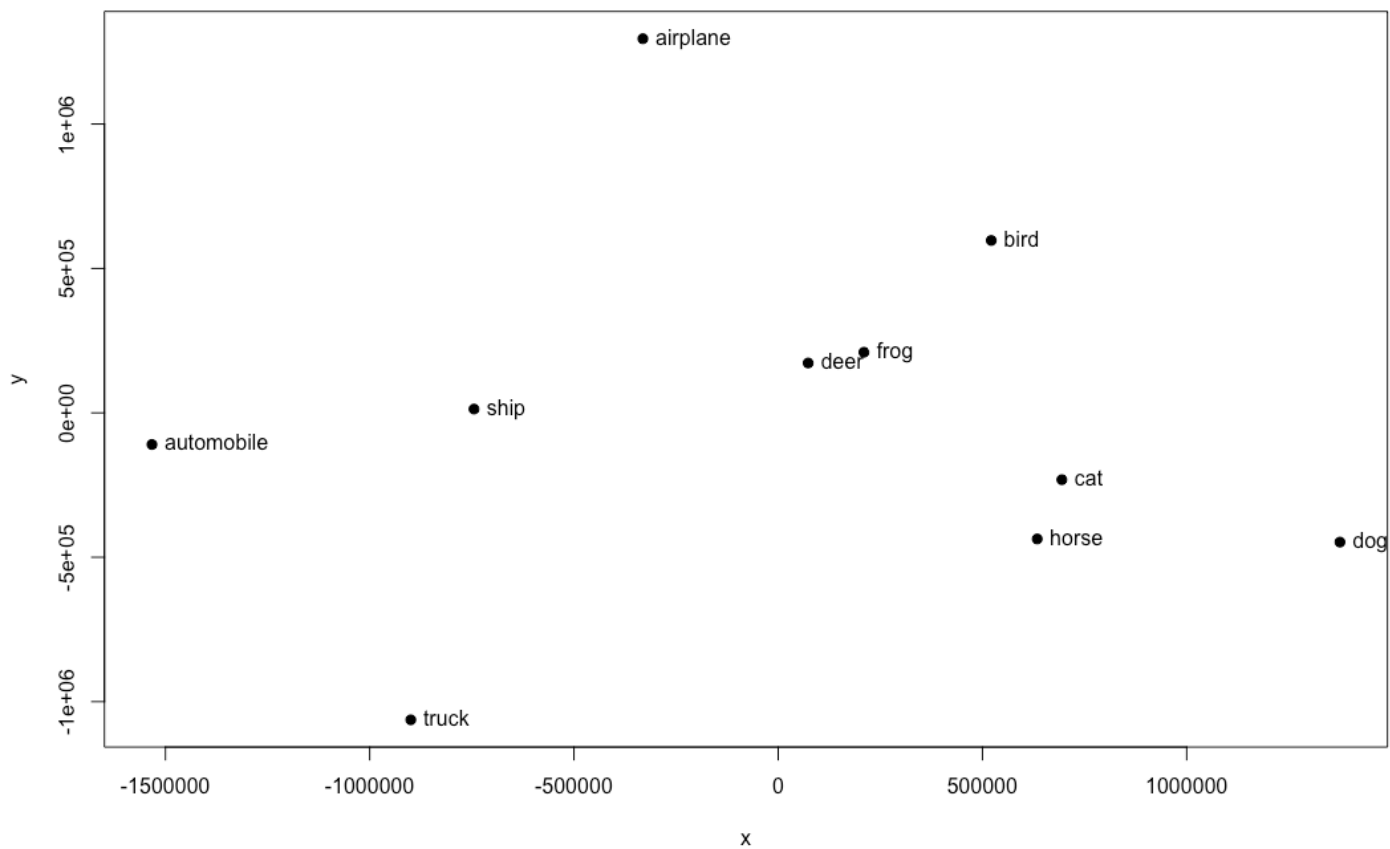
The plot between Part-2 and Part-3 look very similar in terms of how the categories are clustered for most of the images categories and the similarity between them. The low dimensional image reconstruction using the 20 PCs from the same category and the reconstruction using the 20 PCs of other categories in Part 3 have small difference in variance (per pixel difference averaged over all image categories) and that is an interesting fact probably because the first 20 PCs contribute significantly to variance in all image categories & their cumulative variances are similar. The first 20 PCs have drastic impact on change in quality but as more PCs are added there is a sequential improvement in quality.