Brian Cooper (5310799)
coope824@umn.edu

Question 1a:

Model 2: As per equation 5.21 in the Alpaydin textbook, the data can be pooled and a common covariance can be estimated:

$$\boldsymbol{S} = \sum_i P(C_i)S_i$$
$$= \boldsymbol{P(C_1)S_1 + P(C_2)S_2}$$

Model 3:

$$p(x) = \frac{1}{(2\pi)^{d/2}\,|\Sigma|^{1/2}} \exp[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)]$$
$$l(x) = \prod_{t=1}^{N} p(x) = \prod_{t=1}^{N} \frac{1}{(2\pi)^{d/2}\,|\Sigma|^{1/2}} \exp[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)]$$

For $\Sigma = \alpha_1 \mathrm{I}$, maximum likelihood estimation is:

$$\mathcal{L}(x|\mu_1, \alpha_1) = \sum_{t=1}^{N_1} -\frac{d}{2}\log 2\pi - \frac{1}{2}\log|\alpha_1 \mathrm{I}| - \frac{1}{2}(x^t-\mu_1)^T\,(\alpha_1 \mathrm{I})^{-1}(x^t-\mu_1)$$
$$= \sum_{t=1}^{N_1} -\frac{d}{2}\log 2\pi - \frac{1}{2}d\alpha_1 - \frac{1}{2}\left(\frac{1}{\alpha_1}\right)(x^t-\mu_1)^T(x^t-\mu_1)$$

Derivative of MLE with respect to μ_1:

$$\frac{\partial[\mathcal{L}(x|\mu_1,\alpha_1)]}{\partial\mu_1} = \sum_{t=1}^{N_1}(x^t-\mu_1)^T\left(\frac{1}{\alpha_1}\right) = 0$$
$$\Rightarrow \mu_1 = \frac{\sum_{t=1}^{N_1} x^t}{N_1}$$

μ_2 is derived similarly, i.e. $\mu_2 = \frac{\sum_{t=1}^{N_2} x^t}{N_2}$

This result can be converted into a combined μ:

$$\mu = \frac{\sum_{t=1}^{N} r_i^t x^t}{\sum_{t=1}^{N} r_i^t}$$

Derivative of MLE with respect to α_1:

$$\frac{\partial[\mathcal{L}(x|\mu_1,\alpha_1)]}{\partial\alpha_1} = \frac{dN_1}{\alpha_1} - \sum_{t=1}^{N_1}(x^t-\mu_1)^2\left(\frac{1}{\alpha_1^2}\right) = 0$$

$$\Rightarrow d\alpha_1 N_1 - \sum_{t=1}^{N_1}(x^t-\mu_1)^2 = 0$$

$$\Rightarrow \alpha_1 = \frac{\sum_{t=1}^{N_1}(x^t-\mu_1)^2}{N_1 d}$$

α_2 is derived similarly, i.e. $\alpha_2 = \frac{\sum_{t=1}^{N_2}(x^t-\mu_2)^2}{N_2 d}$

This result can be converted into a combined α:

$$\alpha = \frac{\sum_{t=1}^{N} r_i^t\,(x^t-\mu_i)^2}{d\sum_{t=1}^{N} r_i^t}$$

Question 1b, 1c:

Test Set 1, Model 1:
p(C1) = 0.300000

p(C2) = 0.700000

mu1 =
0.430619 2.023520 3.175828 -2.427242 -2.523441 3.237786 -5.520770 -6.692147

mu2 =
4.584063 6.493319 6.426506 1.689060 2.294341 8.362573 -0.165786 -1.804769

S1 =

1.928308	0.234540	0.771907	1.032128	0.432213	1.264847	1.172761	-1.232009
0.234540	3.658970	0.312347	-0.134516	1.582147	1.030130	-0.194275	3.275553
0.771907	0.312347	8.113055	1.334733	-0.428587	1.779151	0.354970	0.232007
1.032128	-0.134516	1.334733	4.229562	0.948213	0.747103	1.068162	1.981016
0.432213	1.582147	-0.428587	0.948213	4.135446	1.002763	-0.545235	3.438464
1.264847	1.030130	1.779151	0.747103	1.002763	4.069565	-0.195951	2.300445
1.172761	-0.194275	0.354970	1.068162	-0.545235	-0.195951	4.216269	-1.709861
-1.232009	3.275553	0.232007	1.981016	3.438464	2.300445	-1.709861	17.095425

S2 =

3.473299	2.099152	2.607964	2.650580	1.798911	1.856803	2.718058	2.976506
2.099152	5.867335	2.210867	2.757547	3.203219	2.929541	2.796325	5.938571
2.607964	2.210867	8.838888	3.424100	2.866566	2.261674	2.793437	5.258646
2.650580	2.757547	3.424100	8.286670	3.629281	2.699248	2.049646	8.518488
1.798911	3.203219	2.866566	3.629281	5.648354	2.948267	3.297118	4.890642
1.856803	2.929541	2.261674	2.699248	2.948267	3.783459	2.267308	4.541467
2.718058	2.796325	2.793437	2.049646	3.297118	2.267308	8.333857	4.367885
2.976506	5.938571	5.258646	8.518488	4.890642	4.541467	4.367885	20.135340

Test set error: 0.200000

Test Set 1, Model 2:

p(C1) = 0.300000

p(C2) = 0.700000

mu1 =

0.430619 2.023520 3.175828 -2.427242 -2.523441 3.237786 -5.520770 -6.692147

mu2 =

4.584063 6.493319 6.426506 1.689060 2.294341 8.362573 -0.165786 -1.804769

S1 =

3.009802	1.539768	2.057147	2.165044	1.388901	1.679216	2.254469	1.713951
1.539768	5.204826	1.641311	1.889928	2.716897	2.359718	1.899145	5.139665
2.057147	1.641311	8.621138	2.797290	1.878020	2.116917	2.061897	3.750654
2.165044	1.889928	2.797290	7.069537	2.824961	2.113605	1.755201	6.557246
1.388901	2.716897	1.878020	2.824961	5.194482	2.364615	2.144412	4.454989
1.679216	2.359718	2.116917	2.113605	2.364615	3.869291	1.528330	3.869161
2.254469	1.899145	2.061897	1.755201	2.144412	1.528330	7.098580	2.544561
1.713951	5.139665	3.750654	6.557246	4.454989	3.869161	2.544561	19.223365

S2 =

3.009802	1.539768	2.057147	2.165044	1.388901	1.679216	2.254469	1.713951
1.539768	5.204826	1.641311	1.889928	2.716897	2.359718	1.899145	5.139665
2.057147	1.641311	8.621138	2.797290	1.878020	2.116917	2.061897	3.750654
2.165044	1.889928	2.797290	7.069537	2.824961	2.113605	1.755201	6.557246
1.388901	2.716897	1.878020	2.824961	5.194482	2.364615	2.144412	4.454989
1.679216	2.359718	2.116917	2.113605	2.364615	3.869291	1.528330	3.869161
2.254469	1.899145	2.061897	1.755201	2.144412	1.528330	7.098580	2.544561
1.713951	5.139665	3.750654	6.557246	4.454989	3.869161	2.544561	19.223365

Test set error: 0.170000

Test Set 1, Model 3:
p(C1) = 0.300000

p(C2) = 0.700000

mu1 =
0.430619 2.023520 3.175828 -2.427242 -2.523441 3.237786 -5.520770 -6.692147

mu2 =
4.584063 6.493319 6.426506 1.689060 2.294341 8.362573 -0.165786 -1.804769

alpha1 = 5.733131

alpha2 = 7.930959

Test set error: 0.240000

Test Set 2, Model 1:
p(C1) = 0.300000

p(C2) = 0.700000

mu1 =
1.065802 2.654770 3.297700 -1.679261 -1.498744 4.395890 -4.213830 -4.967865

mu2 =
2.822076 4.466873 4.853740 0.519233 0.376380 6.258541 -2.661090 -3.817472

S1 =

1.381577	-0.274051	0.683556	0.281881	1.031707	0.478277	-0.297844	0.303881
-0.274051	2.633046	2.136911	0.952171	1.123669	-0.461260	-0.052341	2.784780
0.683556	2.136911	7.151697	1.566701	0.842723	0.199319	2.775914	2.810468
0.281881	0.952171	1.566701	3.157831	0.681055	-0.479989	0.247382	2.380332
1.031707	1.123669	0.842723	0.681055	2.274830	-0.003806	-0.585461	0.774350
0.478277	-0.461260	0.199319	-0.479989	-0.003806	1.246896	0.680311	-0.205408
-0.297844	-0.052341	2.775914	0.247382	-0.585461	0.680311	4.185726	-0.754953
0.303881	2.784780	2.810468	2.380332	0.774350	-0.205408	-0.754953	10.834496

S2 =

2.970742	1.085531	2.951409	0.578655	-0.092368	0.067599	0.886749	1.788748
1.085531	5.062425	2.489295	0.178346	2.771098	-0.328188	0.501476	0.051400
2.951409	2.489295	11.826154	0.089452	0.837207	0.129390	0.022258	-0.845294
0.578655	0.178346	0.089452	5.408230	0.054843	1.024914	4.362731	2.329113
-0.092368	2.771098	0.837207	0.054843	4.316696	1.039629	0.545398	1.692265
0.067599	-0.328188	0.129390	1.024914	1.039629	2.728159	1.731322	2.190835
0.886749	0.501476	0.022258	4.362731	0.545398	1.731322	11.481019	2.751247
1.788748	0.051400	-0.845294	2.329113	1.692265	2.190835	2.751247	13.119085

Test set error: 0.230000

Test Set 2, Model 2:

p(C1) = 0.300000

p(C2) = 0.700000

mu1 =

1.065802	2.654770	3.297700	-1.679261	-1.498744	4.395890	-4.213830	-4.967865

mu2 =

2.822076	4.466873	4.853740	0.519233	0.376380	6.258541	-2.661090	-3.817472

S1 =

2.493993	0.677657	2.271053	0.489623	0.244854	0.190803	0.531371	1.343288
0.677657	4.333611	2.383580	0.410493	2.276869	-0.368110	0.335331	0.871414
2.271053	2.383580	10.423817	0.532627	0.838862	0.150369	0.848355	0.251434
0.489623	0.410493	0.532627	4.733110	0.242706	0.573443	3.128126	2.344479
0.244854	2.276869	0.838862	0.242706	3.704136	0.726599	0.206140	1.416891
0.190803	-0.368110	0.150369	0.573443	0.726599	2.283780	1.416019	1.471963
0.531371	0.335331	0.848355	3.128126	0.206140	1.416019	9.292431	1.699387
1.343288	0.871414	0.251434	2.344479	1.416891	1.471963	1.699387	12.433708

S2 =

2.493993	0.677657	2.271053	0.489623	0.244854	0.190803	0.531371	1.343288
0.677657	4.333611	2.383580	0.410493	2.276869	-0.368110	0.335331	0.871414
2.271053	2.383580	10.423817	0.532627	0.838862	0.150369	0.848355	0.251434
0.489623	0.410493	0.532627	4.733110	0.242706	0.573443	3.128126	2.344479
0.244854	2.276869	0.838862	0.242706	3.704136	0.726599	0.206140	1.416891
0.190803	-0.368110	0.150369	0.573443	0.726599	2.283780	1.416019	1.471963
0.531371	0.335331	0.848355	3.128126	0.206140	1.416019	9.292431	1.699387
1.343288	0.871414	0.251434	2.344479	1.416891	1.471963	1.699387	12.433708

Test set error: 0.560000

Brian Cooper (5310799)
coope824@umn.edu

Test Set 2, Model 3:
p(C1) = 0.300000

p(C2) = 0.700000

mu1 =
1.065802 2.654770 3.297700 -1.679261 -1.498744 4.395890 -4.213830 -4.967865

mu2 =
2.822076 4.466873 4.853740 0.519233 0.376380 6.258541 -2.661090 -3.817472

alpha1 = 3.971320

alpha2 = 7.012434

Test set error: 0.550000
Test Set 3, Model 1:
p(C1) = 0.300000

p(C2) = 0.700000

mu1 =
0.974729 2.623273 3.177043 -1.465240 -1.305277 4.515987 -4.319740 -5.521450

mu2 =
1.491610 3.165544 3.650386 -0.816184 -0.351481 5.134464 -3.277013 -4.729273

S1 =
0.264259 0.084393 0.061156 -0.085321 -0.006019 -0.076364 -0.016044 0.080466
0.084393 0.425812 -0.084981 -0.105143 -0.012187 -0.089623 0.116506 -0.075820
0.061156 -0.084981 0.606287 -0.037661 -0.092520 -0.033265 0.082490 0.013285
-0.085321 -0.105143 -0.037661 0.440020 -0.035577 -0.144641 -0.104437
0.074589
-0.006019 -0.012187 -0.092520 -0.035577 0.432686 0.053367 -0.099753 0.014154
-0.076364 -0.089623 -0.033265 -0.144641 0.053367 0.594422 0.070731 -0.024553
-0.016044 0.116506 0.082490 -0.104437 -0.099753 0.070731 0.544277 -0.030454
0.080466 -0.075820 0.013285 0.074589 0.014154 -0.024553 -0.030454 0.398160

S2 =
2.794477 0.410952 -0.384578 0.598490 -0.412866 0.165961 0.357360 -0.488352
0.410952 2.537093 -0.349723 -0.036702 -0.296160 0.058650 0.178261 0.088163

-0.384578 -0.349723 2.160549 -0.255951 -0.049037 -0.030287 0.035702 -0.283804
0.598490 -0.036702 -0.255951 2.968455 -0.877738 0.102622 -0.342454 -0.327021
-0.412866 -0.296160 -0.049037 -0.877738 3.509480 0.074612 -0.762680 0.194212
0.165961 0.058650 -0.030287 0.102622 0.074612 2.919258 0.642390 0.351849
0.357360 0.178261 0.035702 -0.342454 -0.762680 0.642390 2.947828 0.278364
-0.488352 0.088163 -0.283804 -0.327021 0.194212 0.351849 0.278364 2.612312

Test set error: 0.120000

Test Set 3, Model 2:

p(C1) = 0.300000

p(C2) = 0.700000

mu1 =
0.974729 2.623273 3.177043 -1.465240 -1.305277 4.515987 -4.319740 -5.521450

mu2 =
1.491610 3.165544 3.650386 -0.816184 -0.351481 5.134464 -3.277013 -4.729273

S1 =
2.035411 0.312984 -0.250858 0.393347 -0.290812 0.093264 0.245338 -0.317707
0.312984 1.903709 -0.270301 -0.057234 -0.210968 0.014168 0.159734 0.038968
-0.250858 -0.270301 1.694270 -0.190464 -0.062082 -0.031180 0.049739 -0.194677
0.393347 -0.057234 -0.190464 2.209925 -0.625090 0.028443 -0.271049 -0.206538
-0.290812 -0.210968 -0.062082 -0.625090 2.586442 0.068239 -0.563802 0.140194
0.093264 0.014168 -0.031180 0.028443 0.068239 2.221807 0.470892 0.238928
0.245338 0.159734 0.049739 -0.271049 -0.563802 0.470892 2.226763 0.185719
-0.317707 0.038968 -0.194677 -0.206538 0.140194 0.238928 0.185719 1.948066

S2 =
2.035411 0.312984 -0.250858 0.393347 -0.290812 0.093264 0.245338 -0.317707
0.312984 1.903709 -0.270301 -0.057234 -0.210968 0.014168 0.159734 0.038968
-0.250858 -0.270301 1.694270 -0.190464 -0.062082 -0.031180 0.049739 -0.194677
0.393347 -0.057234 -0.190464 2.209925 -0.625090 0.028443 -0.271049 -0.206538
-0.290812 -0.210968 -0.062082 -0.625090 2.586442 0.068239 -0.563802 0.140194
0.093264 0.014168 -0.031180 0.028443 0.068239 2.221807 0.470892 0.238928
0.245338 0.159734 0.049739 -0.271049 -0.563802 0.470892 2.226763 0.185719
-0.317707 0.038968 -0.194677 -0.206538 0.140194 0.238928 0.185719 1.948066

Brian Cooper (5310799)
coope824@umn.edu

Test set error: 0.450000

Test Set 3, Model 3:

p(C1) = 0.300000

p(C2) = 0.700000

mu1 =
0.974729 2.623273 3.177043 -1.465240 -1.305277 4.515987 -4.319740 -5.521450

mu2 =
1.491610 3.165544 3.650386 -0.816184 -0.351481 5.134464 -3.277013 -4.729273

alpha1 = 0.447799

alpha2 = 2.766093

Test set error: 0.050000

Test set performance from best to worst

Set 1:	2, 1, 3
Set 2:	1,3,2
Set 3:	3,1,2

Model 1 works best (lowest test error) on data set 2, model 2 works best on data set 1, and model 3 works best on data set 3. This is due to different S_1 and S_2 values working better on different distributions of data.

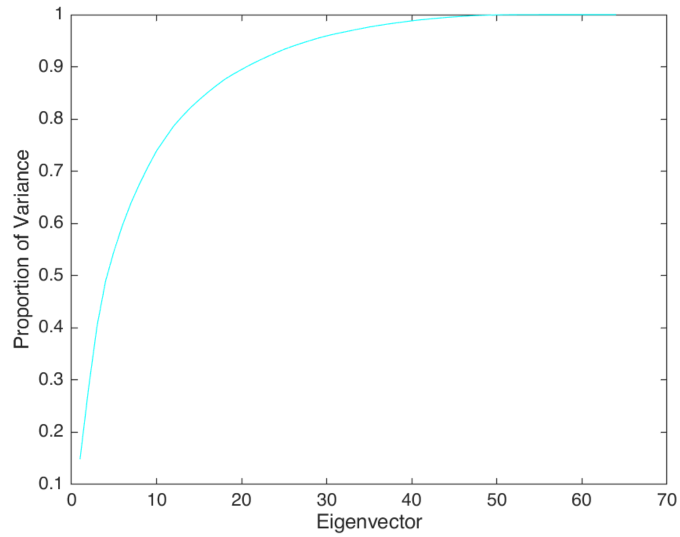
Question 2a:

MATLAB Output:
Error rate for k = 1: 0.053872
Error rate for k = 3: 0.040404
Error rate for k = 5: 0.043771
Error rate for k = 7: 0.053872

Brian Cooper (5310799)
coope824@umn.edu

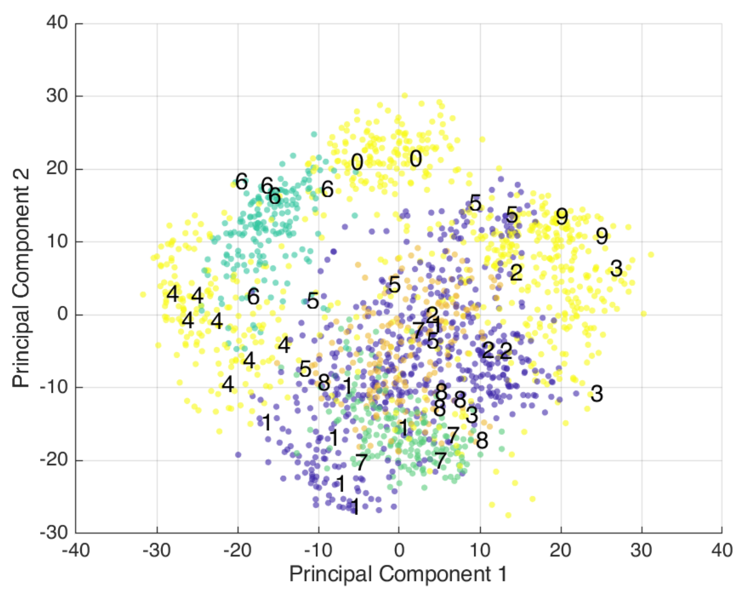
Question 2b:

MATLAB Output:



Minimum K value that explains at least 90% of variance: 21
Testing error rate for k = 1: 0.043771
Testing error rate for k = 3: 0.040404
Testing error rate for k = 5: 0.043771
Testing error rate for k = 7: 0.040404

Question 2c:



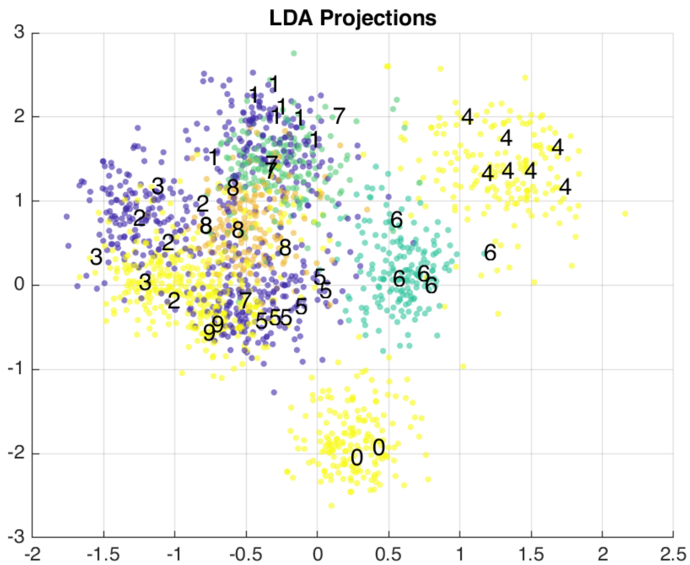
Brian Cooper (5310799)
coope824@umn.edu

Question 2d:

MATLAB Output:

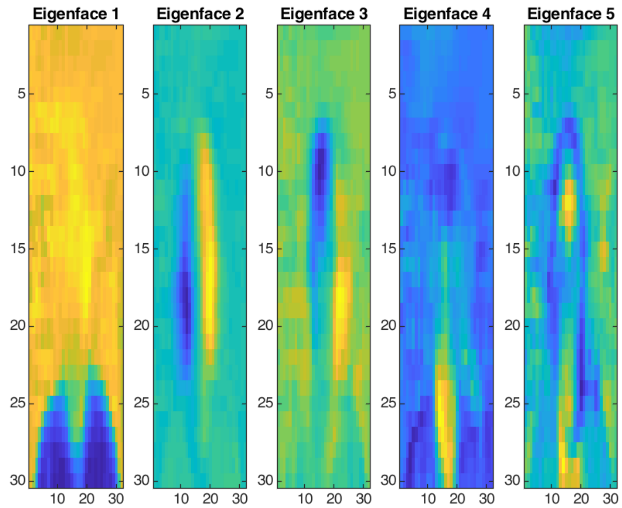
Error rate for (L = 2), (k = 1): 0.464646
Error rate for (L = 2), (k = 3): 0.424242
Error rate for (L = 2), (k = 5): 0.407407
Error rate for (L = 4), (k = 1): 0.191919
Error rate for (L = 4), (k = 3): 0.181818
Error rate for (L = 4), (k = 5): 0.158249
Error rate for (L = 9), (k = 1): 0.097643
Error rate for (L = 9), (k = 3): 0.094276
Error rate for (L = 9), (k = 5): 0.090909

Question 2e:



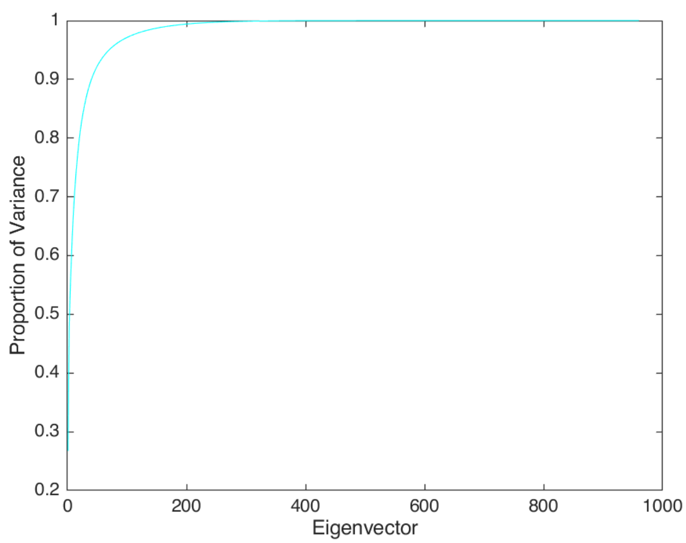
Brian Cooper (5310799)
coope824@umn.edu

Question 3a:



Question 3b:

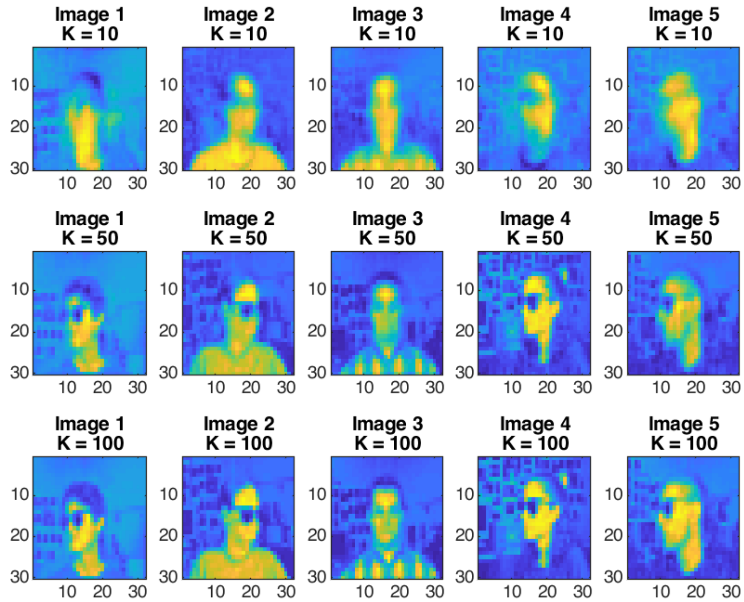
MATLAB Output:



Minimum K value that explains at least 90% of variance: 41
Testing error rate for k = 1: 0.104839
Testing error rate for k = 3: 0.241935
Testing error rate for k = 5: 0.395161
Testing error rate for k = 7: 0.395161

Brian Cooper (5310799)
coope824@umn.edu

Question 3c:



With a lower value for K, the reconstructed images appear to be blurrier. Increasing K appears to sharpen the images.