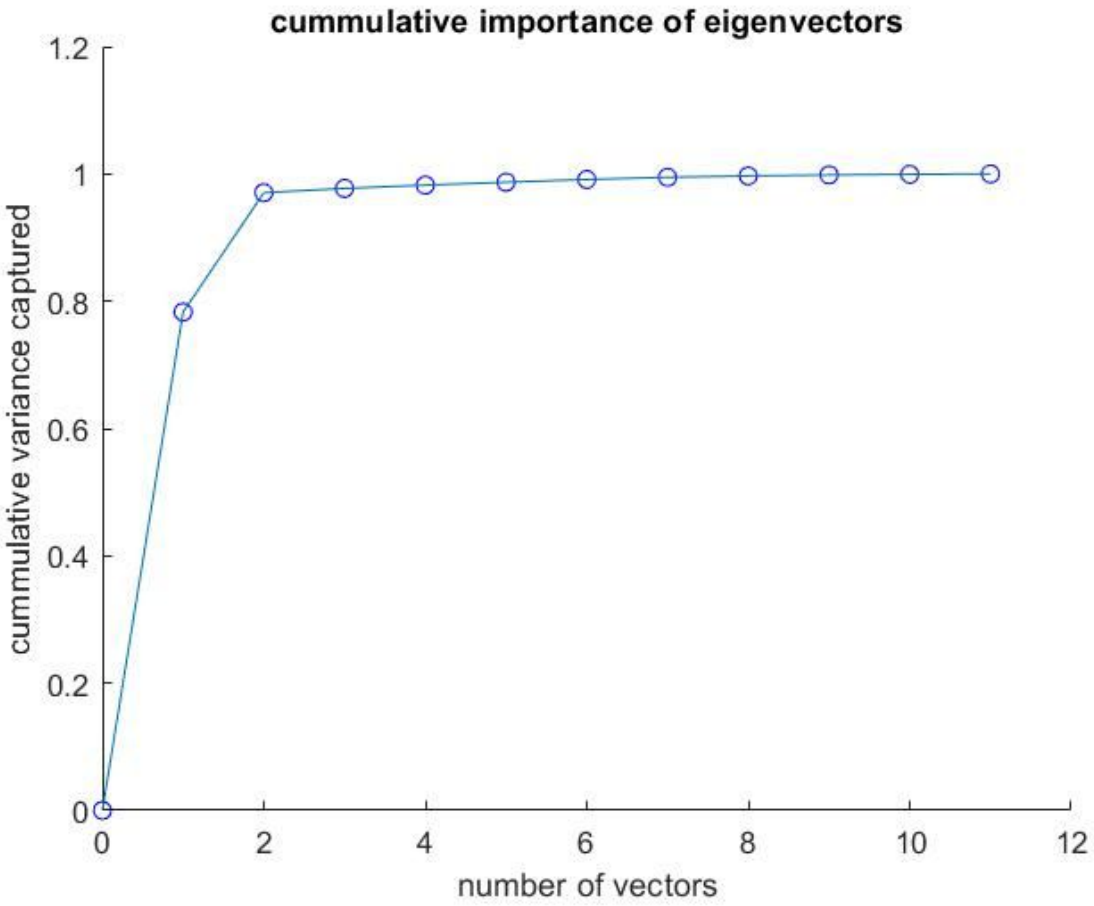
1. 10 visits are used to get a more thorough understanding of what each person buys, rather than possibly getting just “the one day this person bought a pickle and nothing else”.

5.

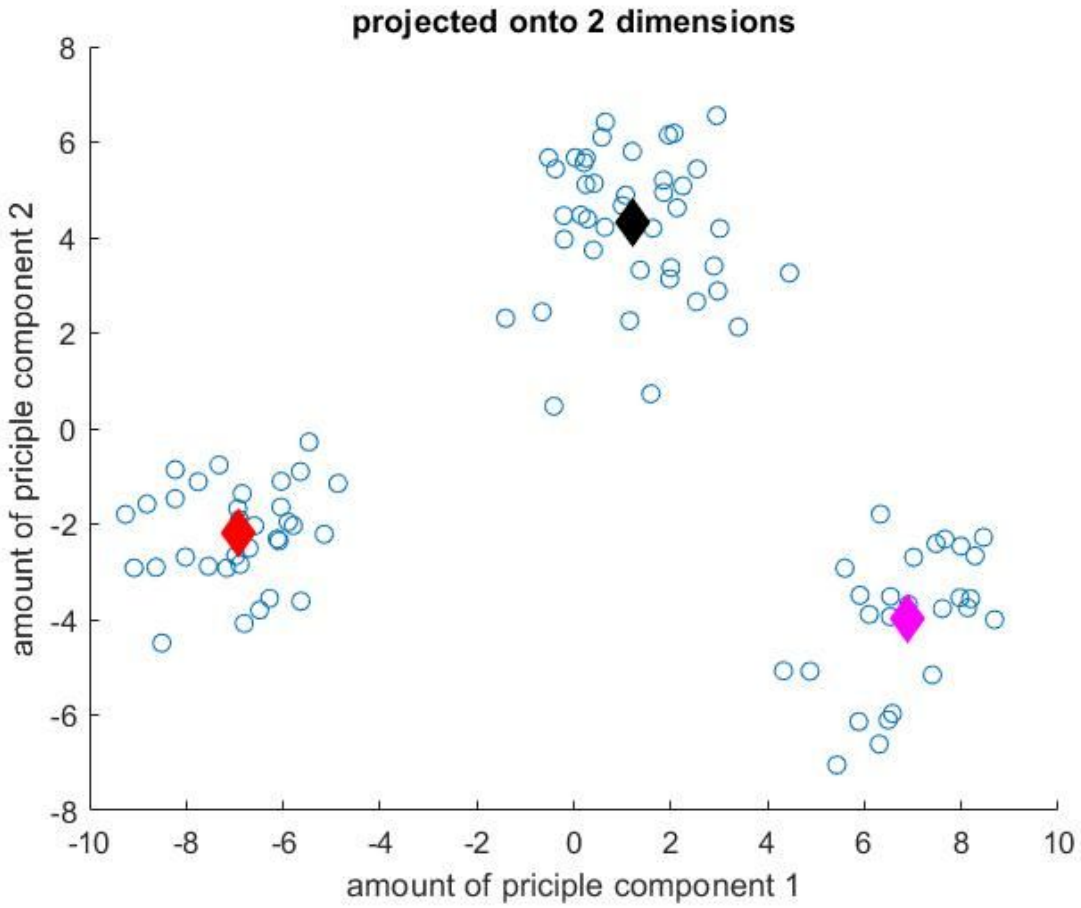


6. the two most useful eigenvectors are:

0.2749	0.5703
-0.0021	-0.0096
-0.3095	0.0599
0.1576	0.2800
-0.0202	-0.4056
-0.3427	0.3418
0.3675	0.0929
0.3126	0.3317
-0.3550	-0.0480
0.2068	-0.2359
0.4071	-0.3517
-0.3444	0.1056

Because the second value in each vector, which corresponds to petfood, are very near to zero, petfood can be ignored. I did not ignore it however because this time it was easier not to remove it as removing it would require that covariance and pca be done all over again to obtain eigenvectors of the appropriate sizes.

7.



9. 2d space cluster centers:

-6.9292 -2.1946
6.8723 -3.9960
1.2191 4.3004

10. Back to full 12d space:

	1	2	3	4	5	6	7	8	9	10	11	12
1	-3.1565	0.0355	2.0131	-1.7067	1.0300	1.6247	-2.7506	-2.8936	2.5656	-0.9153	-2.0488	2.1548
2	-0.3894	0.0241	-2.3664	-0.0357	1.4818	-3.7210	2.1546	0.8227	-2.2480	2.3641	4.2026	-2.7889
3	2.7875	-0.0438	-0.1197	1.3963	-1.7688	1.0520	0.8476	1.8073	-0.6394	-0.7624	-1.0160	0.0342

These values seem quite reasonable once considering that they are deviations from the mean of all records we began with, including petfood, which all 3 clusters have an average of near zero here, which means the population average is also almost the average for each cluster.

11. During this assignment i learned a bit about how the matlab pca, covariance, and k-means functions worked by looking up the documentation pages when i needed them for the program.