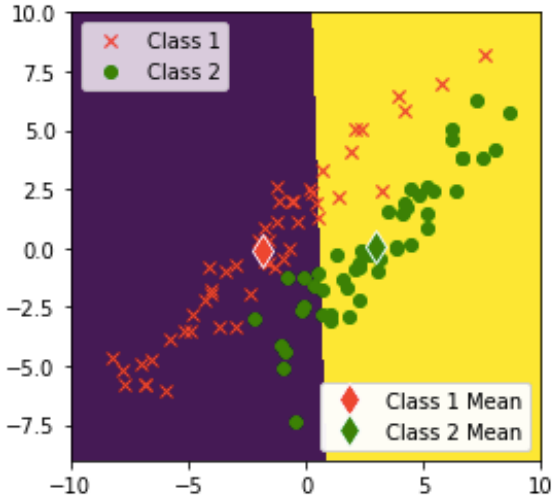
(a)

synthetic1

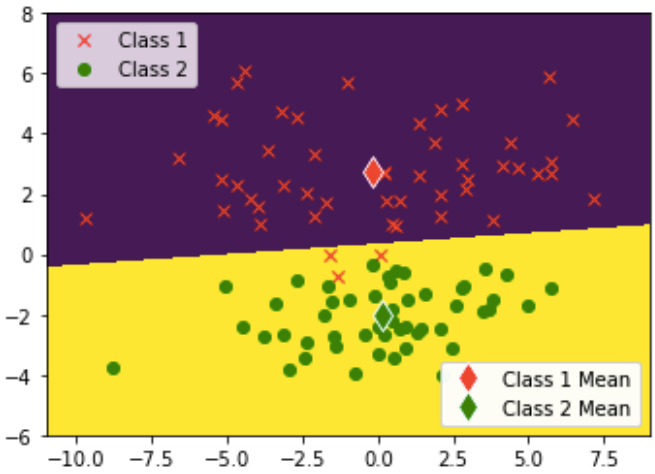
```
class 0 mean is : [-1.8731152 -0.1166418]
class 1 mean is : [2.98095798 0.03548129]
```



error rate for training data is 21.0%
error rate for testing data is 24.0%

synthetic2

```
class 0 mean is : [-0.2032685   2.75522592]
class 1 mean is : [ 0.13275594 -2.0526066 ]
```



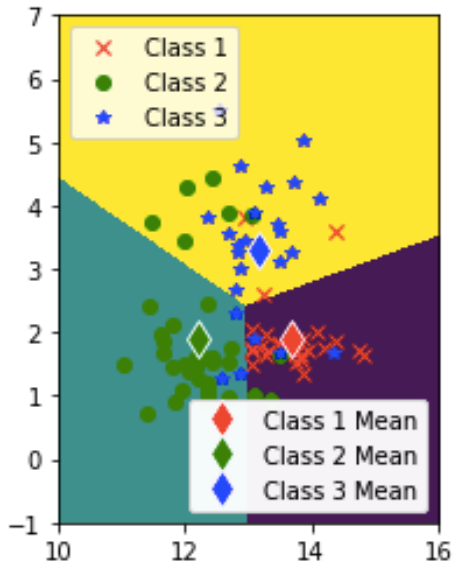
error rate for training data is 3.0%
error rate for testing data is 4.0%

(b)

Yes, there is much difference.
Because data points in synthetic1 is more disperse and higher variance. Also from the plot we can know that data points are not around the meant point, they lie on a line. So if we use the distance to mean point to classify the data points, this method won't work well for this kind of data.
However, for synthetic2, all the data points spread evenly around mean point. So for most of the data, the distance to mean point is much smaller than that to another class.

(c)

```
class 0 mean is : [13.675   1.904]
class 1 mean is : [12.21457143  1.88885714]
class 2 mean is : [13.17333333  3.28333333]
```



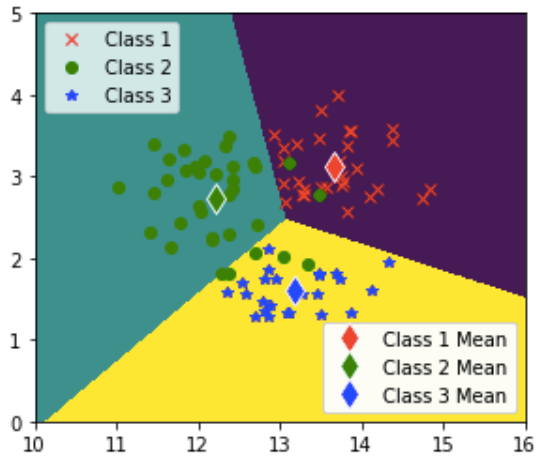
error rate for wine train data is 20.2247191011%
error rate for wine test data is 22.4719101124%

(d)

In order to find the pair data to minimize the error rate, I went through all the possible pairs and calculate the error rate for each pair. For example [0,1], [0,2]...[0,12]…[11,12] Then I chose the pair that have minimum error rate.

According to the result, The index of the pair that minimize the error rate is [0,11] which is the first column and the 12^{th} column.

```
The minimum error_rate is: 7.86516853933%
Which means the error number is: 7
And the index is: [0, 11]
class 0 mean is : [13.675  3.127]
class 1 mean is : [12.21457143  2.73714286]
class 2 mean is : [13.17333333  1.59416667]
```



error rate for training data is 7.86516853933%
error rate for test data is 12.3595505618%

(e)

Yes, I think there is much difference in training-set error rate for different pairs of features. As it shows on the left side, for different sets, the error rate changes from over 44% to 8.9%

```
[0, 1]
20.2247191011%
[0, 2]
31.4606741573%
[0, 3]
44.9438202247%
[0, 4]
56.1797752809%
[0, 5]
14.606741573%
[0, 6]
8.98876404494%
[0, 7]
33.7078651685%
[0, 8]
16.8539325843%
[0, 9]
25.8426966292%
```

```
[0, 1]
22.4719101124%
[0, 2]
26.9662921348%
[0, 3]
41.5730337079%
[0, 4]
38.202247191%
[0, 5]
15.7303370787%
[0, 6]
8.98876404494%
[0, 7]
28.0898876404%
[0, 8]
23.5955056618%
[0, 9]
23.5955056618%
```

And the same thing happens for testing data set, the table on the right side is the error rate for test data, which could change from over 41% to 8.98%.