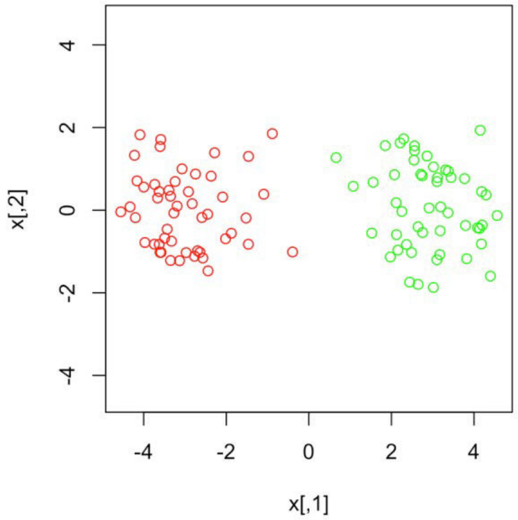
Consider two point clouds ($n=100$ each), randomly drawn around two origins 3 units away from the origin:

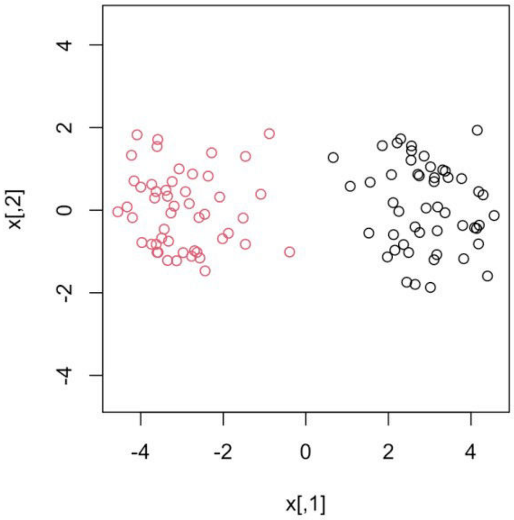
```
set.seed(495)
n <- 100
d <- 3
x <- matrix(rnorm(n * 2, sd = 1), ncol = 2)
x[1:(n/2), 1] <- x[1:(n/2), 1] - d
x[(n/2 + 1):n, 1] <- x[(n/2 + 1):n, 1] + d
```



The K-means algorithm has no problem in classifying these points:

```
km <- kmeans(x, centers = 2)
km$centers
```

```
##        [,1]         [,2]
## 1  2.922143  0.098422541
## 2 -2.991026 -0.003131757
```

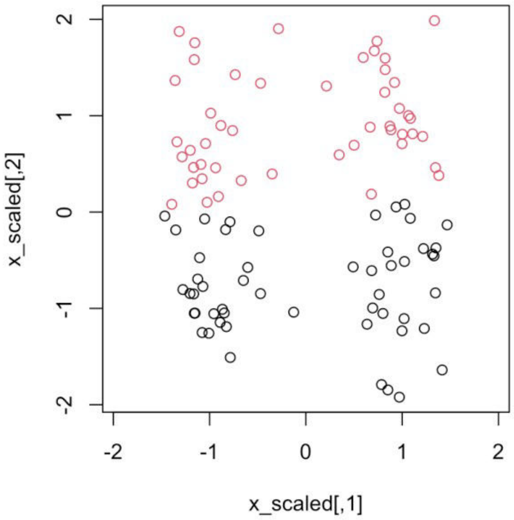


Let’s see now what happens when we standardize each feature. Since their mean is already zero, we merely divide by their standard deviation:

```
x_scaled <- x
x_scaled[, 1] <- x_scaled[, 1] / sd(x_scaled[, 1])
x_scaled[, 2] <- x_scaled[, 2] / sd(x_scaled[, 2])
```

And we run again the K-means algorithm on these new data:

```
km_scaled <- kmeans(x_scaled, centers = 2)
```



We see that K-means has completely failed to identify the clusters, because ‘standardizing’ the features has destroyed the clear separation between the clusters.