

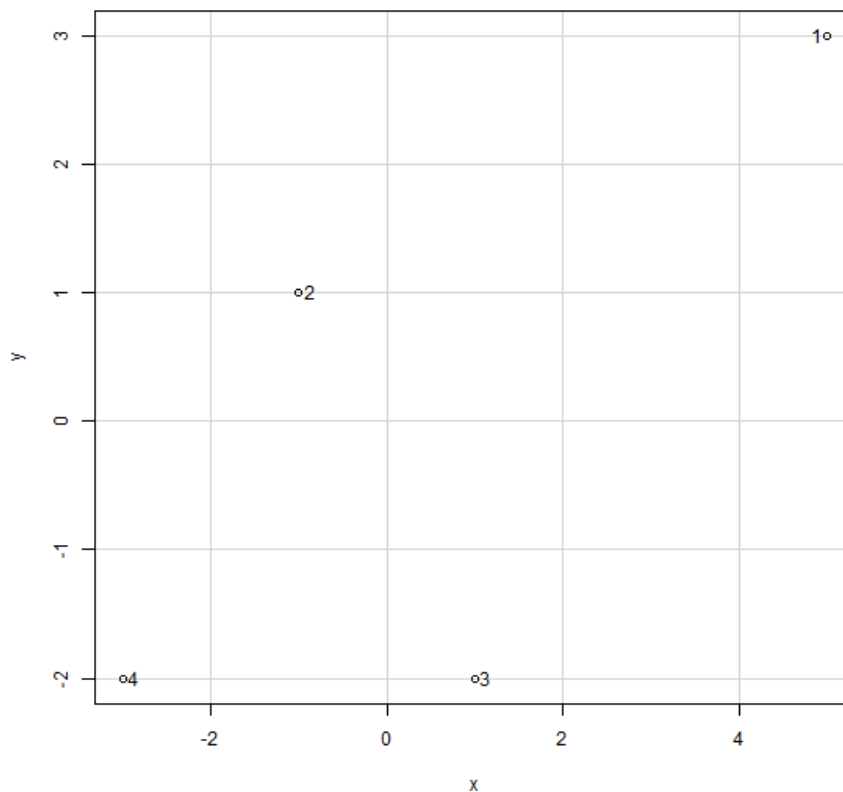
Manual k-means

MP

2015-06-23

```
> Dataset <-
+ read.table("D:/Dropbox/R/2015-NUS/Session-3/(b) k-Means/Manual/Manual-k.csv",
+ header=TRUE, sep="," , na.strings="NA", dec=".", strip.white=TRUE)
```

```
> scatterplot(y~x, reg.line=FALSE, smooth=FALSE, spread=FALSE,
+ id.method='mahal', id.n = 5, boxplots=FALSE, span=0.5, data=Dataset)
```



```
1 2 3 4
1 2 3 4
```

```
> .cluster <- KMeans(model.matrix(~-1 + x + y, Dataset), centers = 2,
+ iter.max = 10, num.seeds = 10)
> .cluster$size # Cluster Sizes
```

```
[1] 3 1
```

```
> .cluster$centers # Cluster Centroids
```

```
new.x.x new.x.y
1      -1      -1
2       5       3
```

```
> .cluster$withinss # Within Cluster Sum of Squares
```

```
[1] 14 0
```

```
> .cluster$tot.withinss # Total Within Sum of Squares
```

```
[1] 14
```

```
> .cluster$betweenss # Between Cluster Sum of Squares
```

```
[1] 39
```

```
> Dataset$KMeans <- assignCluster(model.matrix(~-1 + x + y, Dataset), Dataset,  
+   .cluster$cluster)  
> remove(.cluster)
```