

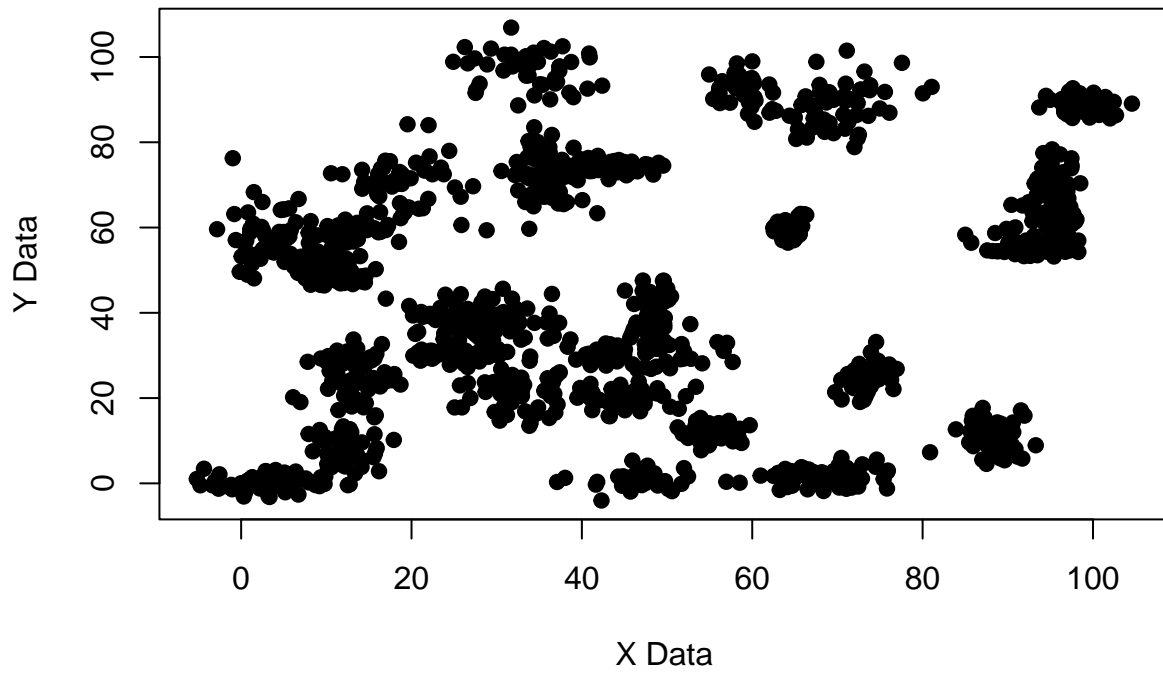
# Assignment 10

Joshua Greenert

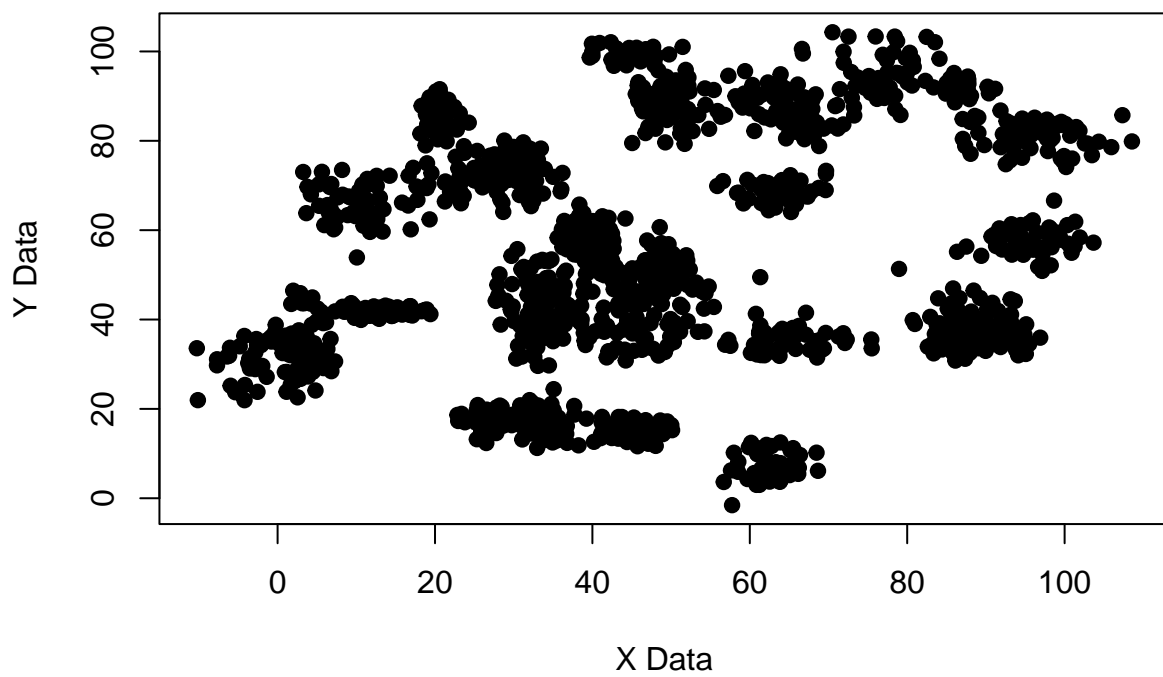
2022-08-12

## Scatterplots

### Binary Classification Scatterplot



### Trinary Classification Scatterplot



## Looking at the plots, would linear classification work here?

Since the data appears to be scattered all over the x and y axis, it's unlikely that a linear classification would be effective since the models aren't linear.

## Binary K Nearest Neighbors

```
## [1] "When k equals 3, accuracy is: 0.98"
## [1] "When k equals 5, accuracy is: 0.978"
## [1] "When k equals 10, accuracy is: 0.978"
## [1] "When k equals 15, accuracy is: 0.9719"
## [1] "When k equals 20, accuracy is: 0.98"
## [1] "When k equals 25, accuracy is: 0.9739"
```

## Trinary K Nearest Neighbors

```
## [1] "When k equals 3, accuracy is: 0.9273"
## [1] "When k equals 5, accuracy is: 0.9197"
## [1] "When k equals 10, accuracy is: 0.9272"
## [1] "When k equals 15, accuracy is: 0.8738"
## [1] "When k equals 20, accuracy is: 0.8891"
## [1] "When k equals 25, accuracy is: 0.8891"
```

## Accuracy Results for Binary and Trinary

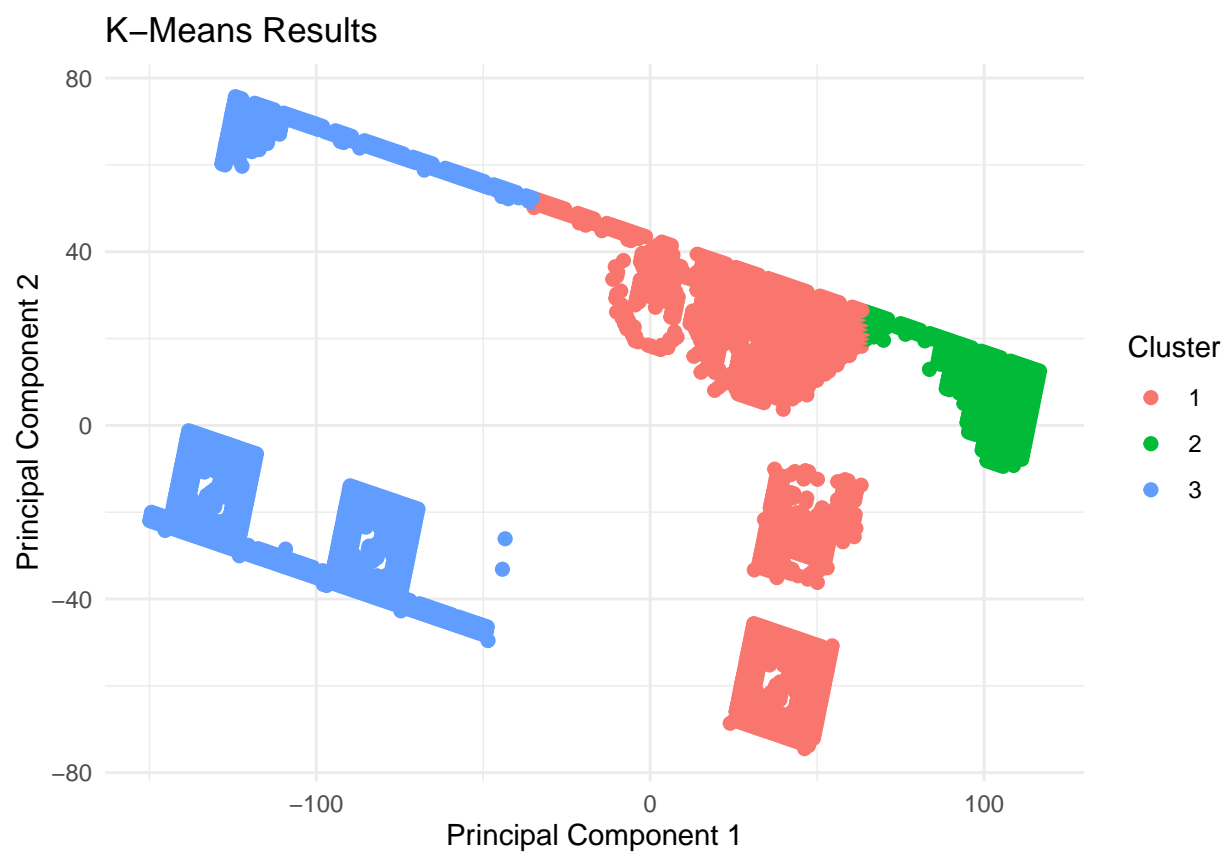
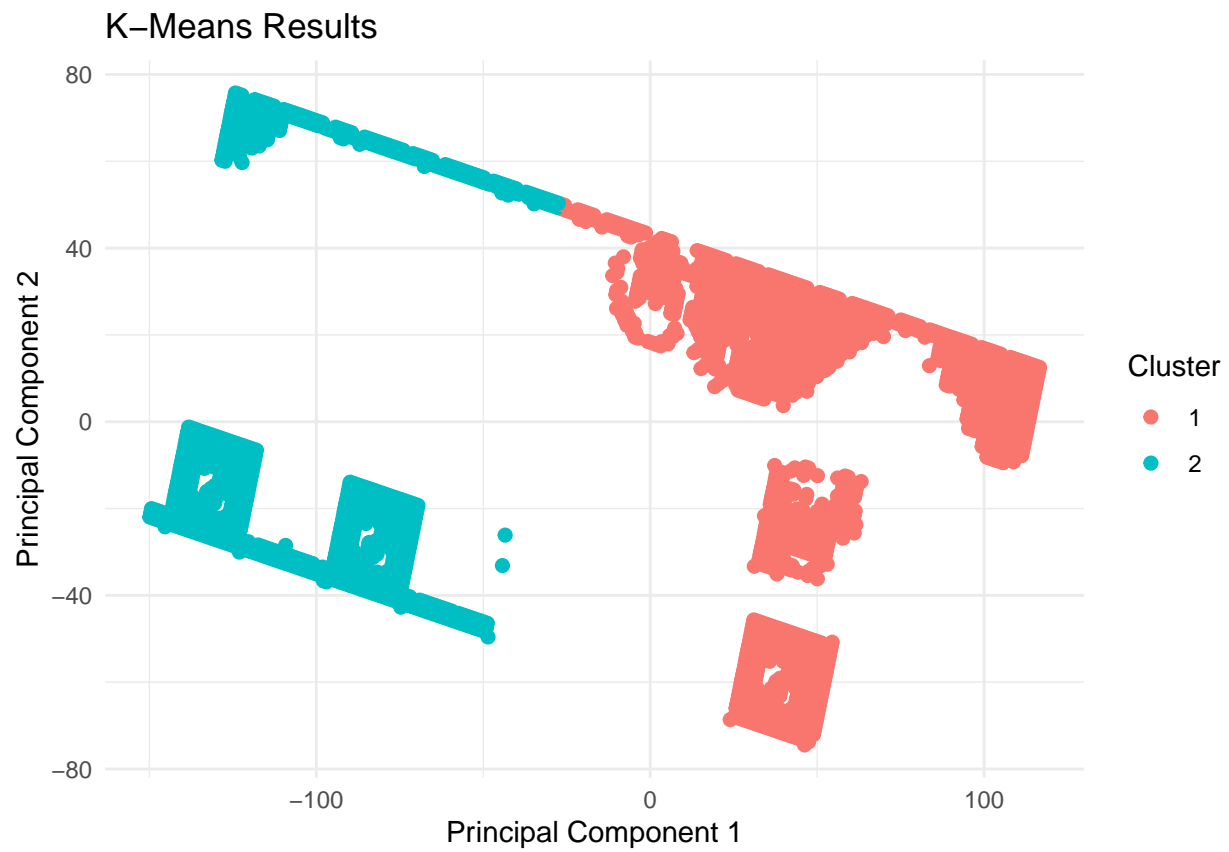
With the binary and the trinary datasets, we can see that our highest degree of accuracy happens when  $k = 3$  as the binary dataset shows 98.00% and the trinary dataset shows us 94.83%. As we increase  $k$ , we see that the accuracy percentage reduces.

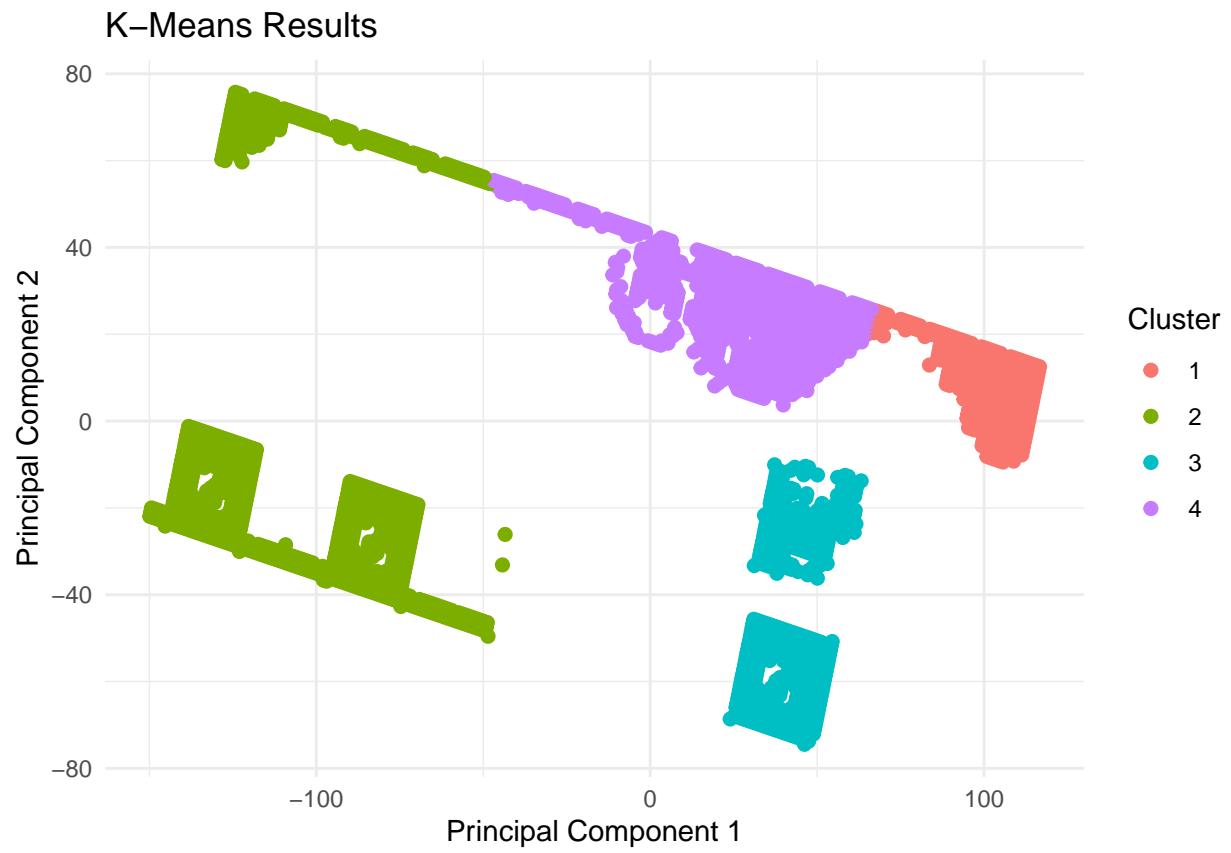
## Clustering

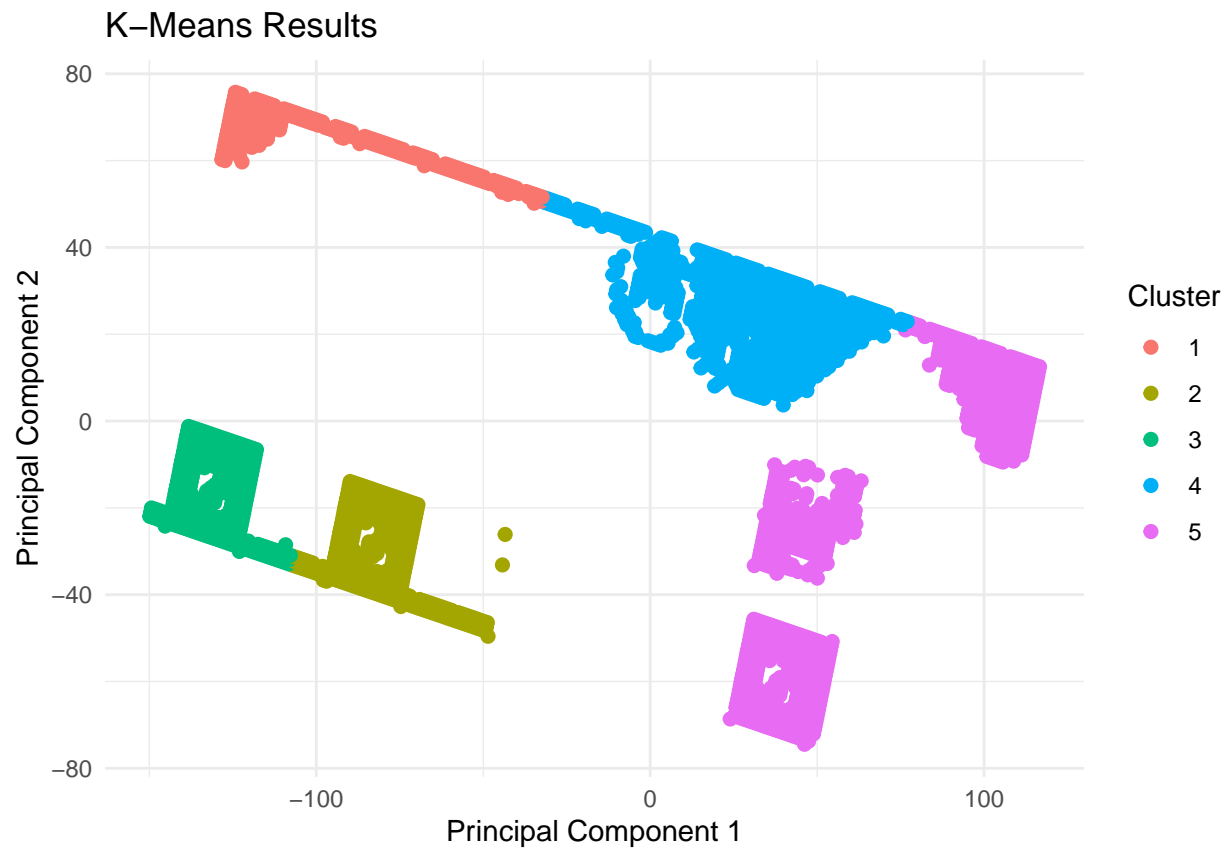
### Cluster Distance Average

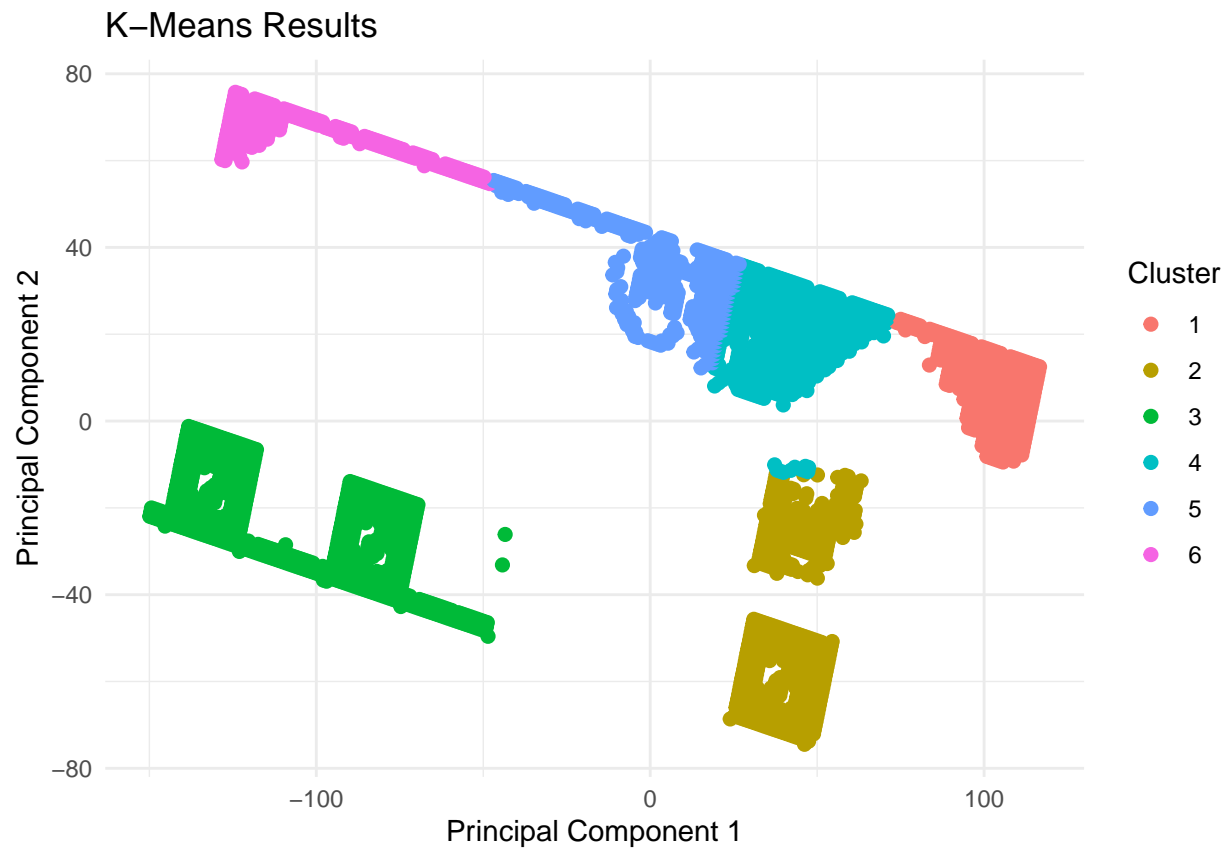
```
## [1] 101.4671
```

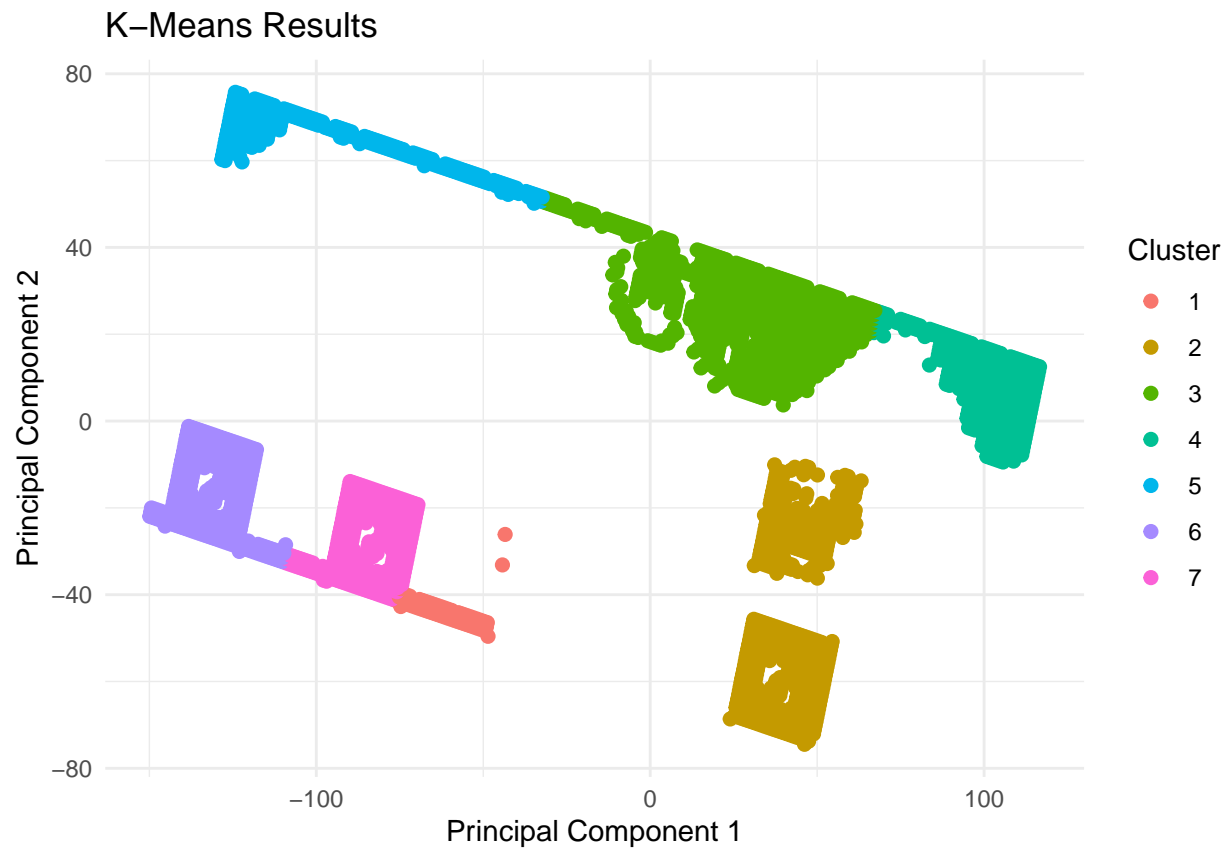
## Cluster K-Means



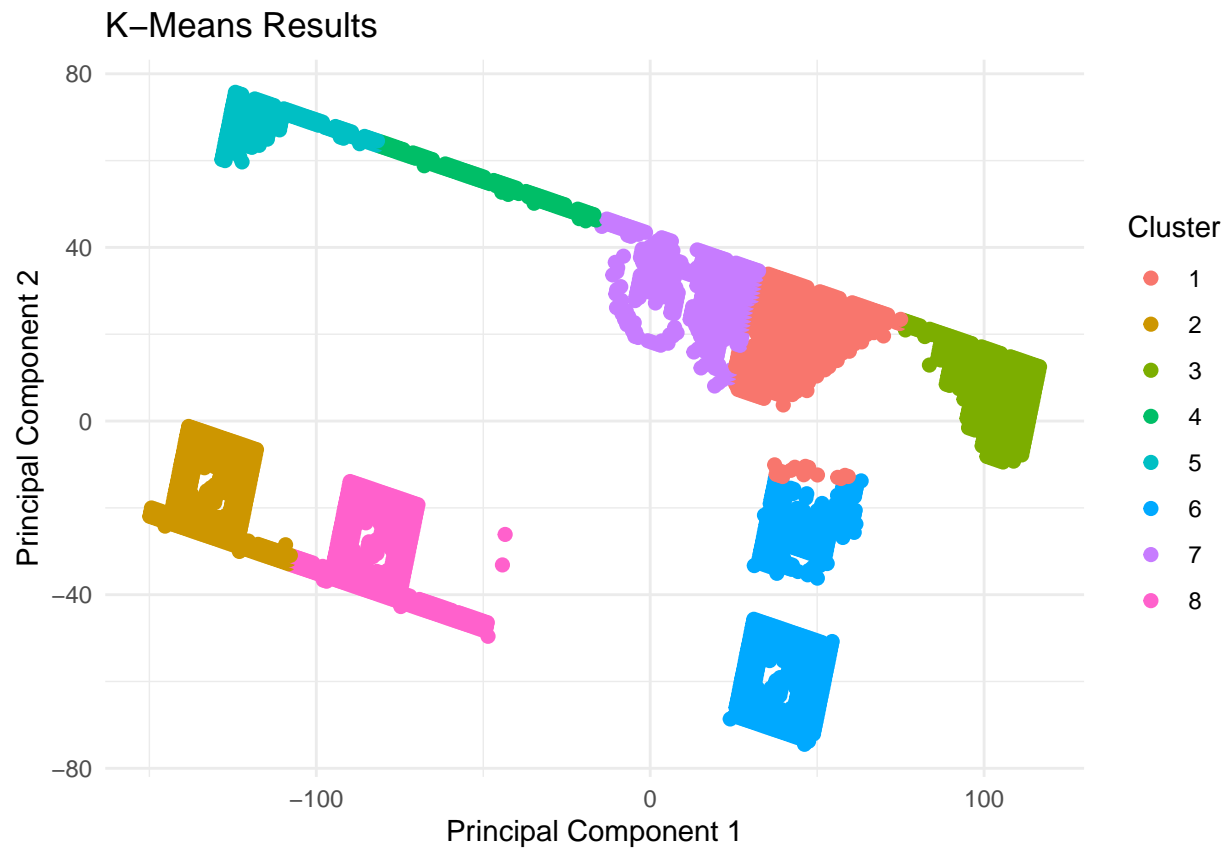


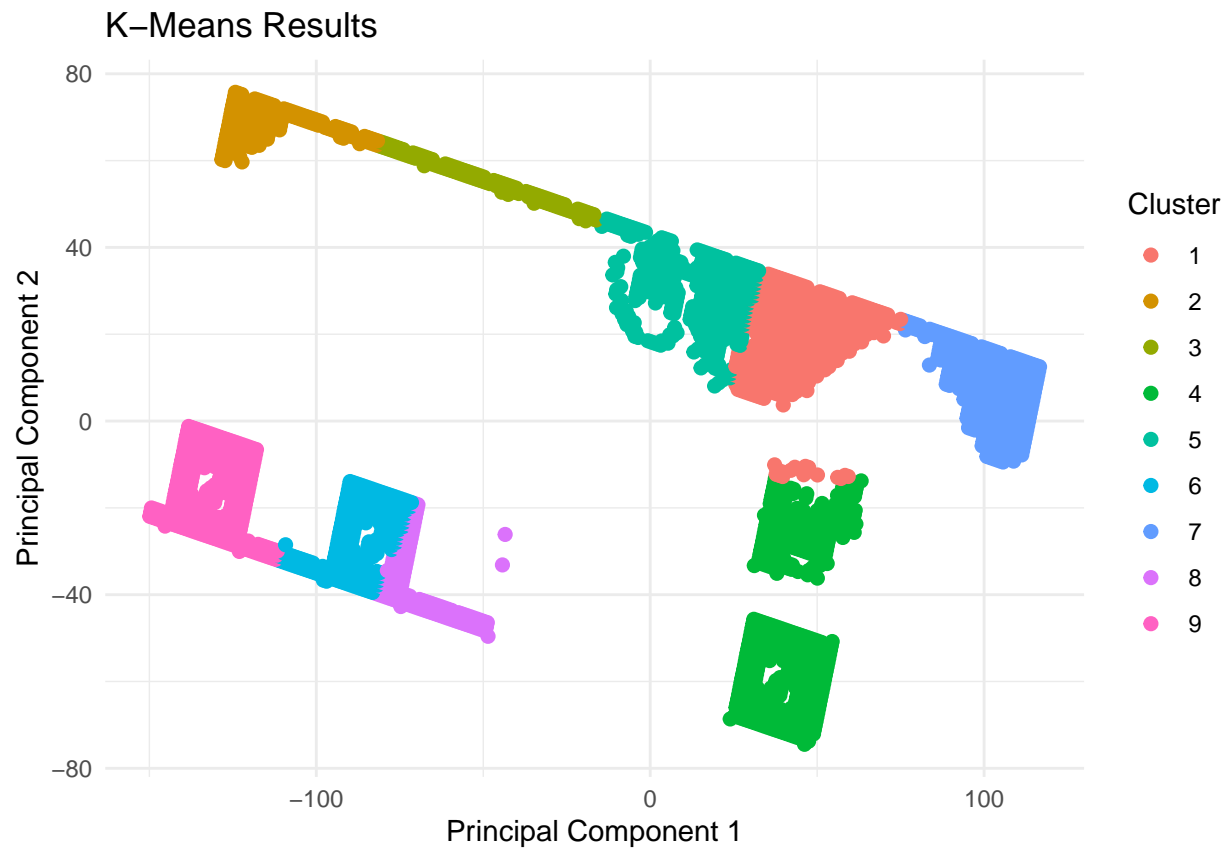


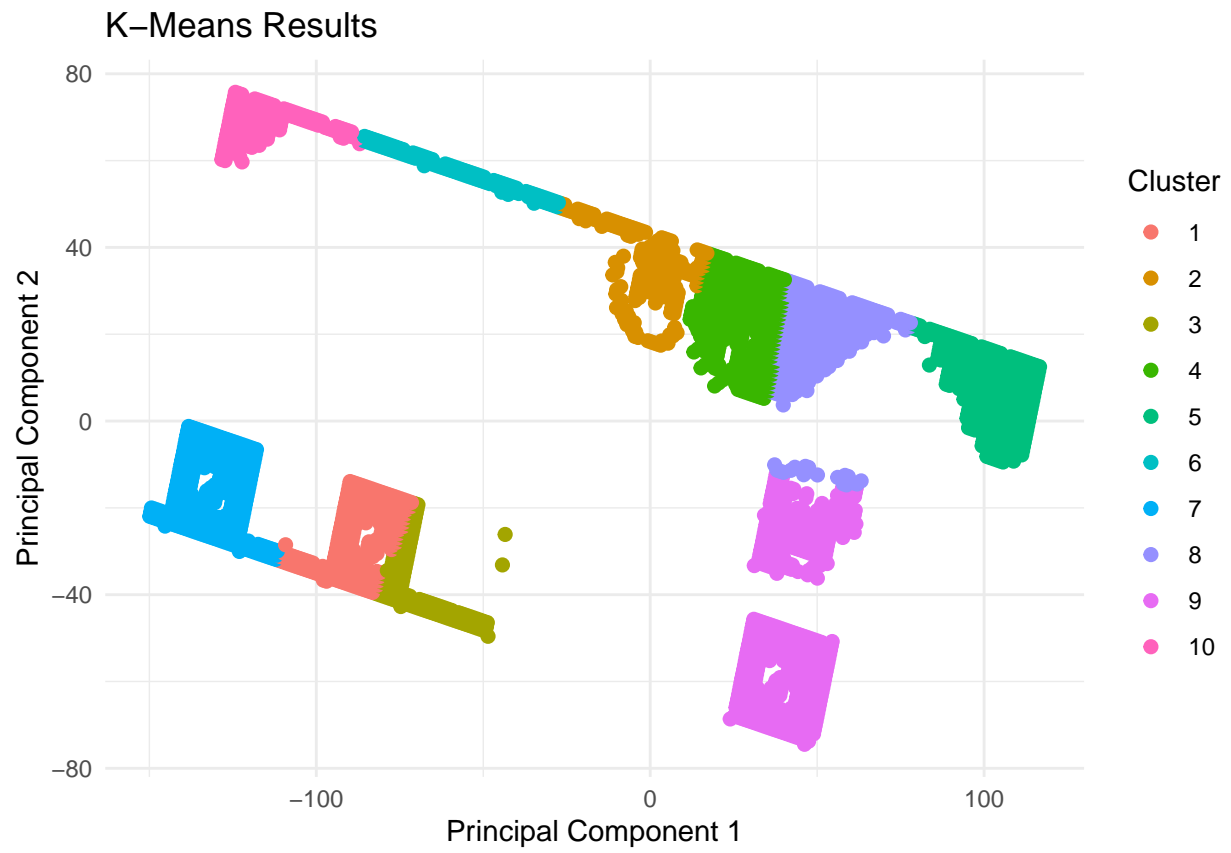


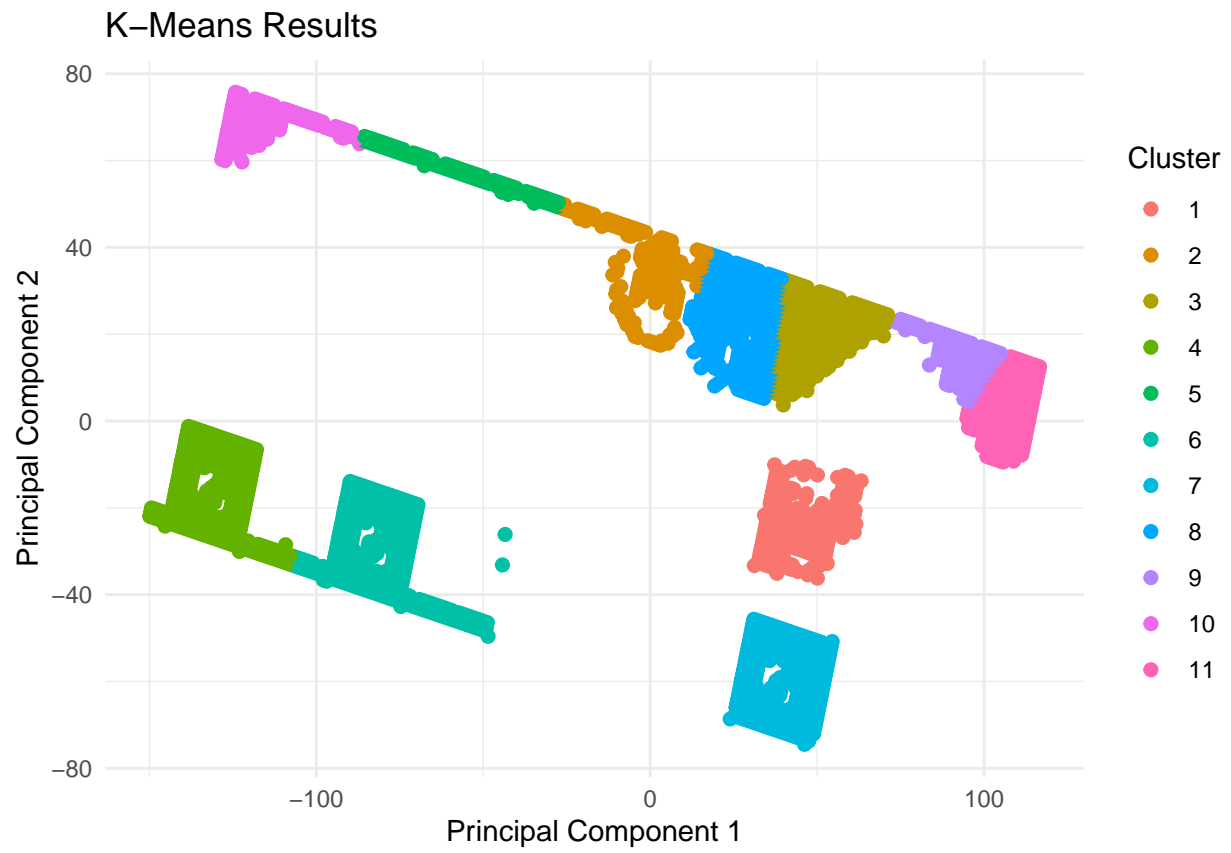


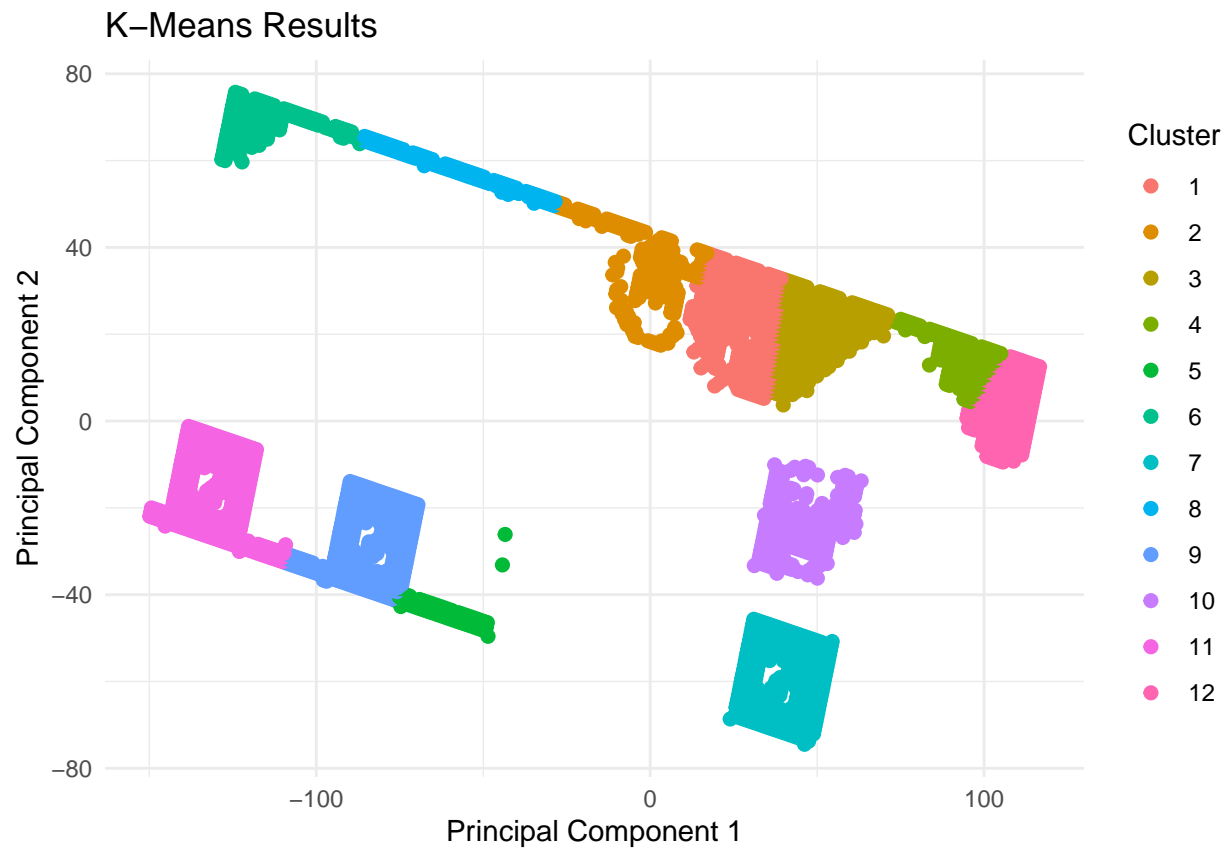












## Cluster Elbow Points

