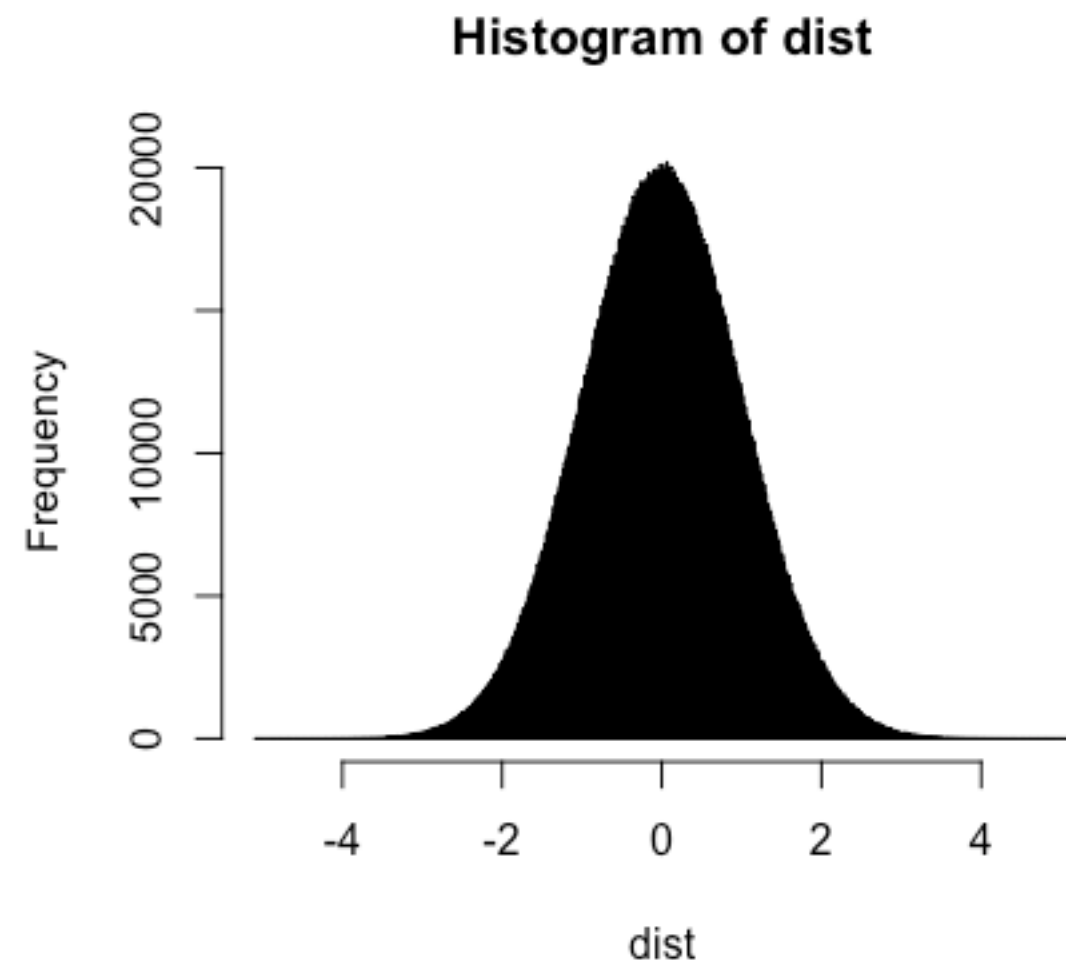


- The larger the sample the more representative it is of the population it is drawn from. Here is a $N = 5$ million sample from the standard normal distribution.

```
> set.seed(1234)
```

```
> dist <- rnorm(5000000, 0, 1)
```

```
> hist(dist, breaks = 1000)
```



- In the standard normal distribution, 95% of the data is contained within 1.96 standard deviations either side of the mean.

