Lab-02

Basic Stats

1. Data Sampling

Here we create a vector of 20 random values

```
x <- c(10, 2, 15, 6, 4, 9, 12, 11, 3, 0, 12, 10, 9, 7, 11, 10, 8, 5, 10, 6)

x
```

```
##  [1] 10  2 15  6  4  9 12 11  3  0 12 10  9  7 11 10  8  5 10  6
```

Here we get the number of elements in the vector

```
length(x)
```

```
## [1] 20
```

This command sums the numbers in the vector

```
sum(x)
```

```
## [1] 160
```

This command gives the mean of the numbers in the vector

```
mean(x)
```

```
## [1] 8
```

This one gives the median

```
median(x)
```

```
## [1] 9
```

This command computes the mode of the vector

```
names(sort(-table(x)))[1]
```

```
## [1] "10"
```

Here we extract the Quantile Ranges from the vector (Min, Q1, M, Q3, Max)

```
quantile(x)
```

```
##    0%   25%   50%   75%  100%
##  0.00  5.75  9.00 10.25 15.00
```

Here we compute the variance

```
var(x)
```

```
## [1] 14.52632
```

This command computes the standard deviation

```
sd(x)
```

```
## [1] 3.81134
```

2. Distrbutions

Here we return a new vector y, with a normal distribution, with 20 observations, the same mean as vector x, and a standard deviation of 1

```
y <- rnorm(20, mean(x), 1)

y
```

```
##  [1] 9.433189 7.113756 9.555344 8.851715 8.457588 7.858684 7.550448
##  [8] 7.425489 8.313684 7.629427 8.324027 6.874065 7.425408 9.840942
## [15] 8.243414 8.332950 8.033881 7.776898 7.750533 8.998992
```

This command computes the co-varaince of x and y

```
cov(x, y)
```

```
## [1] 0.4983055
```

This command computes the correlation of x and y

```
cor(x, y)
```

```
## [1] 0.1603206
```

Because we have not set a seed value, the value of y will change everytime y is call. Hence the values of the co-varaince & correlation between x & y will also change everytime. We therefore set a seed value, re-assign value to y, with the same parameters

```
set.seed(63)

y <- rnorm(20, mean(x), 1)

y
```

```
##  [1] 9.324112 6.126815 8.492406 6.100057 8.547002 6.928197 8.402325
##  [8] 8.132398 6.346650 8.423790 8.382545 7.856132 8.340911 8.178827
## [15] 7.757927 8.217518 7.613421 8.912965 8.954544 9.191293
```

3. Scatter Plots

We now create 2 normal distributions, each with 20 observations, and a standard deviation of 1. Vector z has a mean of 8

```
set.seed(100)

z <- rnorm(20, 8, 1)

z
```

```
##  [1]  7.497808  8.131531  7.921083  8.886785  8.116971  8.318630  7.418209
##  [8]  8.714533  7.174741  7.640138  8.089886  8.096274  7.798366  8.739840
## [15]  8.123380  7.970683  7.611146  8.510856  7.086186 10.310297
```

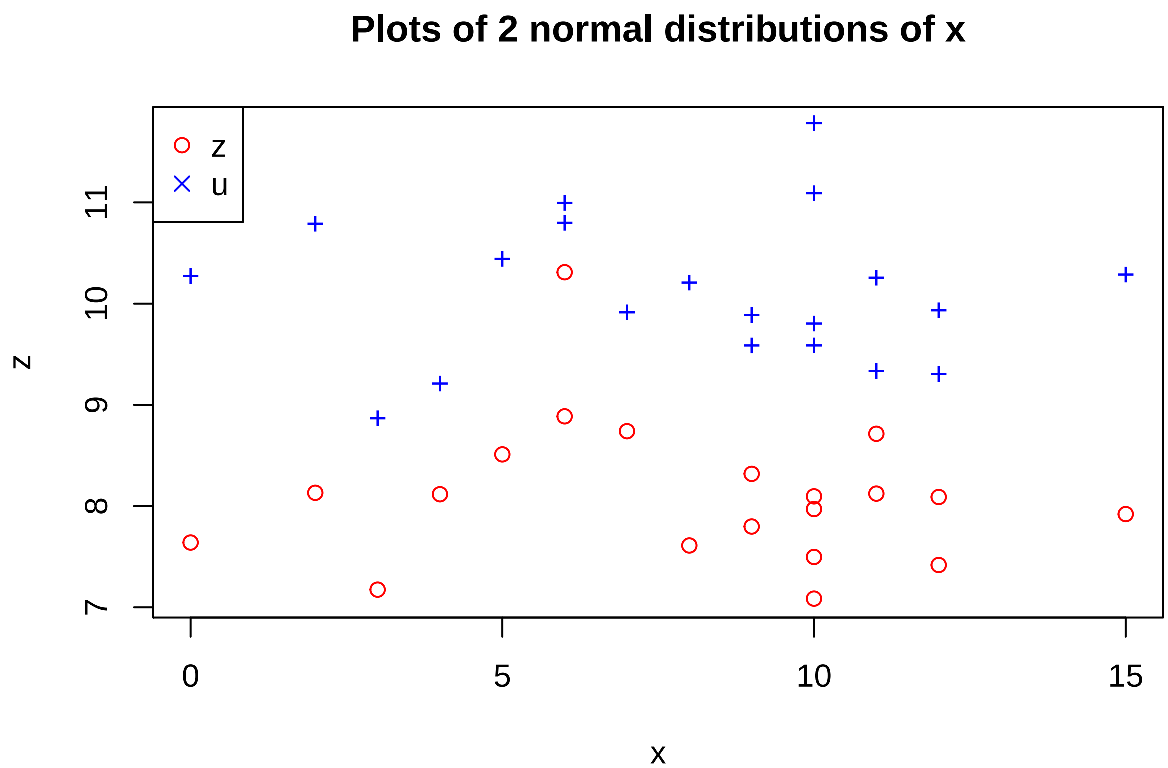
Vector u has a mean of 10

```
u <- rnorm(20, 10, 1)

u
```

```
## [1]  9.561910 10.764061 10.261961 10.773405  9.185621  9.561549  9.279778
## [8] 10.230945  8.842271 10.247076  9.908886 11.757376  9.862070  9.888807
## [15]  9.309986  9.778206 10.182908 10.417323 11.065402 10.970202
```

This command creates a scatter plots of (x,z) & (x,u) & a legend

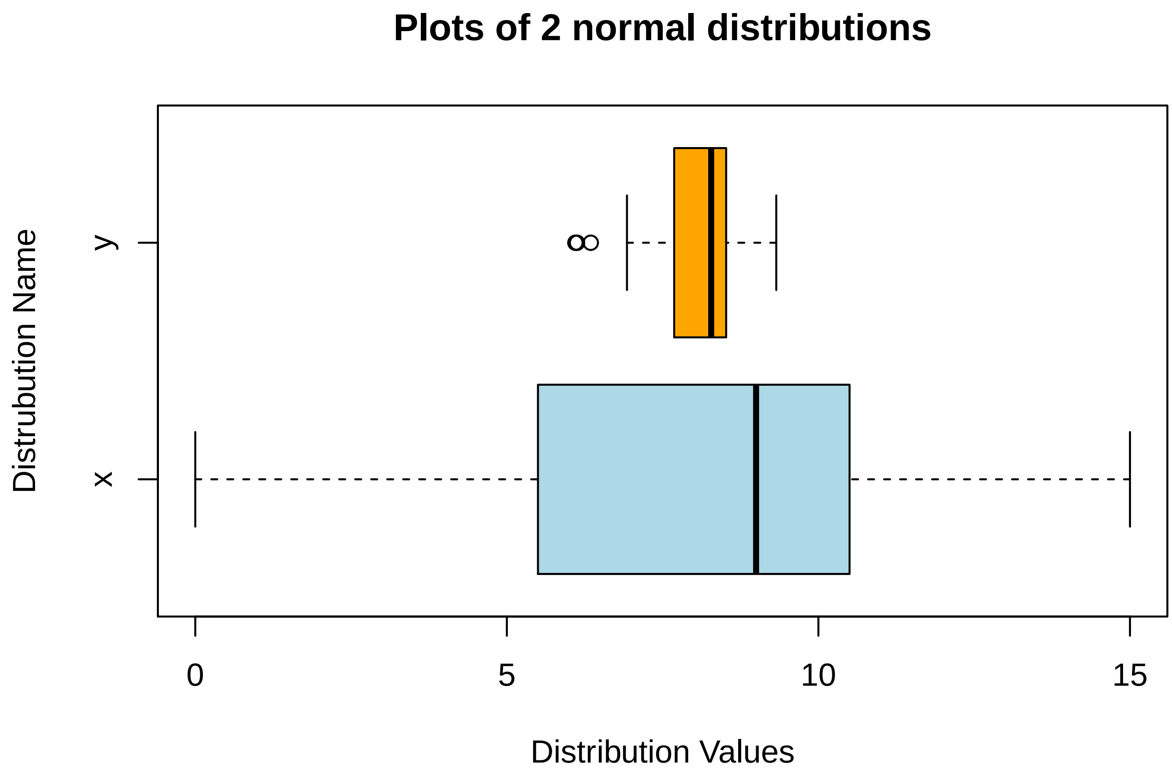
```r
plot(x,z, xlim=range(x, x), ylim=range(z, u), col="red", main= "Plots of 2 normal distributions of x")
points(x,u, pch='+', col='blue')

legend("topleft",pch=c(1,4), col=c("red", "blue"), c("z", "u"), box.col="black")
```



4. Box Plots

Here we are creating box plots of x and y, against the same axis

```r
boxplot(x,y, names =(c("x", "y")), col=(c("lightblue","orange")),
        main="Plots of 2 normal distributions", xlab="Distribution Values", ylab="Distrubution Name",
        horizontal=TRUE)
```


Plots of 2 normal distributions
Distrubution Name
y
x
0
5
10
15
Distribution Values