

RomilGupta_LAB_RDS.R

Romil

Fri Nov 03 00:02:08 2017

```
# Exercise 1
ds1<- read.csv(file="E:/Data Science/Assignments/R/29-Oct-17/ds1.csv" ,header=TRUE,sep=",")
ds2<- read.csv(file="E:/Data Science/Assignments/R/29-Oct-17/ds2.csv" ,header=TRUE,sep=",")
ds3<- read.csv(file="E:/Data Science/Assignments/R/29-Oct-17/ds3.csv" ,header=TRUE,sep=",")
ds4<- read.csv(file="E:/Data Science/Assignments/R/29-Oct-17/ds4.csv" ,header=TRUE,sep=",")
#summary of all the Dataset
summary(ds1)
```

```
##           X1           Y2
## Min.      : 4.0      Min.      : 4.260
## 1st Qu.: 6.5      1st Qu.: 6.315
## Median : 9.0      Median : 7.580
## Mean     : 9.0      Mean      : 7.546
## 3rd Qu.:11.5      3rd Qu.: 8.820
## Max.     :14.0      Max.      :10.840
```

```
summary(ds2)
```

```
##           X2           Y2
## Min.      : 4.0      Min.      :3.100
## 1st Qu.: 6.5      1st Qu.:6.695
## Median : 9.0      Median :8.140
## Mean     : 9.0      Mean      :7.501
## 3rd Qu.:11.5      3rd Qu.:8.950
## Max.     :14.0      Max.      :9.260
```

```
summary(ds3)
```

```
##           X3           Y3
## Min.      : 4.0      Min.      : 5.39
## 1st Qu.: 6.5      1st Qu.: 6.25
## Median : 9.0      Median : 7.11
## Mean     : 9.0      Mean      : 7.50
## 3rd Qu.:11.5      3rd Qu.: 7.98
## Max.     :14.0      Max.      :12.74
```

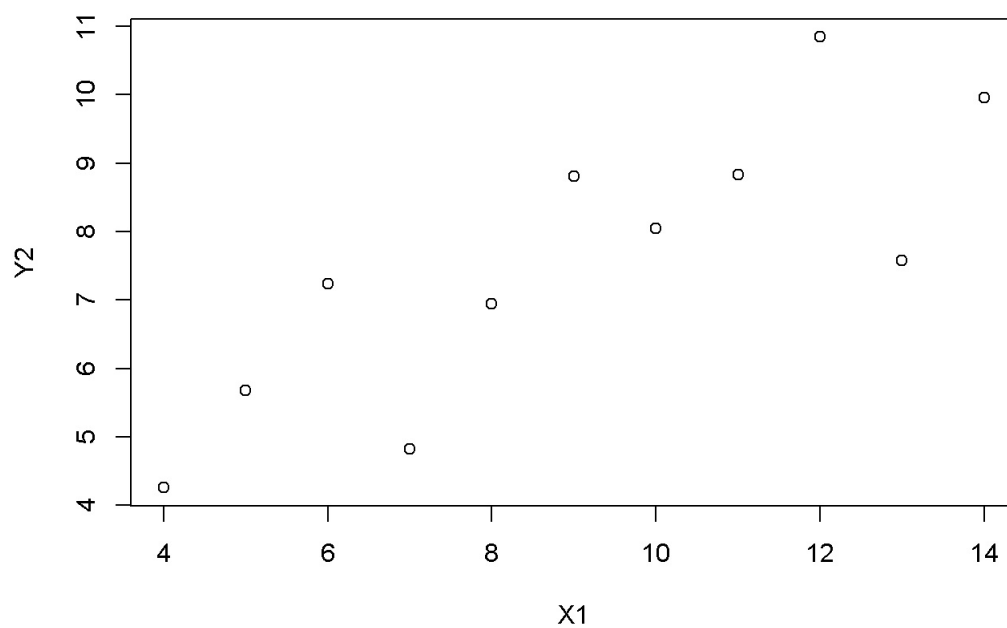
```
summary(ds4)
```

```
##           X4           Y4
## Min.      : 8      Min.      : 5.250
## 1st Qu.: 8      1st Qu.: 6.170
## Median : 8      Median : 7.040
## Mean     : 9      Mean      : 7.501
## 3rd Qu.: 8      3rd Qu.: 8.190
## Max.     :19      Max.      :12.500
```

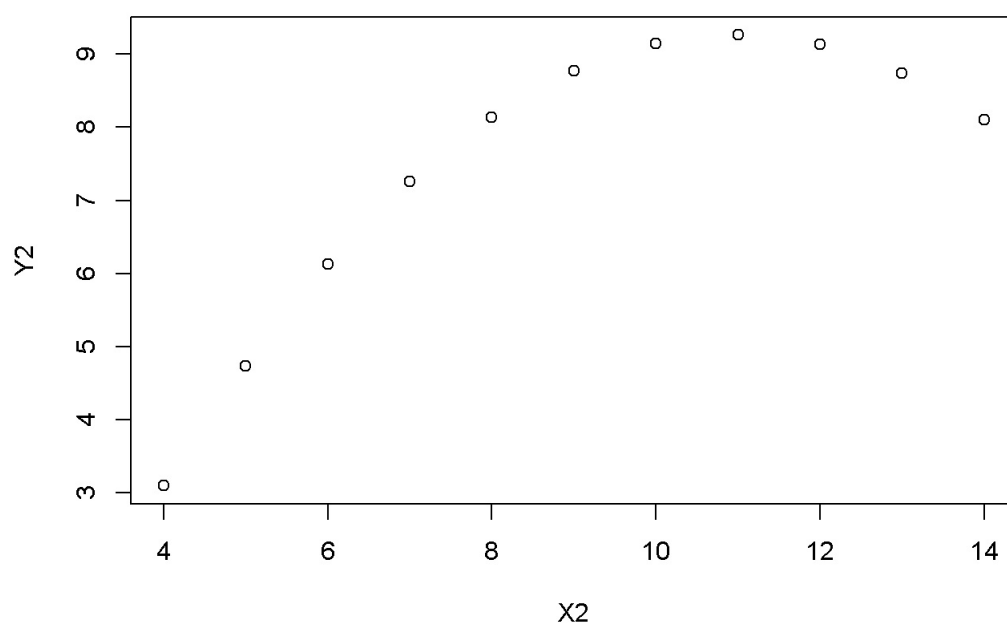
```
# Exercise 2
daf<- data.frame(ds1,ds2,ds3,ds4)
summary(daf)
```

```
##           X1           Y2           X2           Y2.1
## Min.      : 4.0      Min.      : 4.260      Min.      : 4.0      Min.      :3.100
## 1st Qu.: 6.5      1st Qu.: 6.315      1st Qu.: 6.5      1st Qu.:6.695
## Median : 9.0      Median : 7.580      Median : 9.0      Median :8.140
## Mean      : 9.0      Mean      : 7.546      Mean      : 9.0      Mean      :7.501
## 3rd Qu.:11.5      3rd Qu.: 8.820      3rd Qu.:11.5      3rd Qu.:8.950
## Max.      :14.0      Max.      :10.840      Max.      :14.0      Max.      :9.260
##           X3           Y3           X4           Y4
## Min.      : 4.0      Min.      : 5.39      Min.      : 8      Min.      : 5.250
## 1st Qu.: 6.5      1st Qu.: 6.25      1st Qu.: 8      1st Qu.: 6.170
## Median : 9.0      Median : 7.11      Median : 8      Median : 7.040
## Mean      : 9.0      Mean      : 7.50      Mean      : 9      Mean      : 7.501
## 3rd Qu.:11.5      3rd Qu.: 7.98      3rd Qu.: 8      3rd Qu.: 8.190
## Max.      :14.0      Max.      :12.74      Max.      :19      Max.      :12.500
```

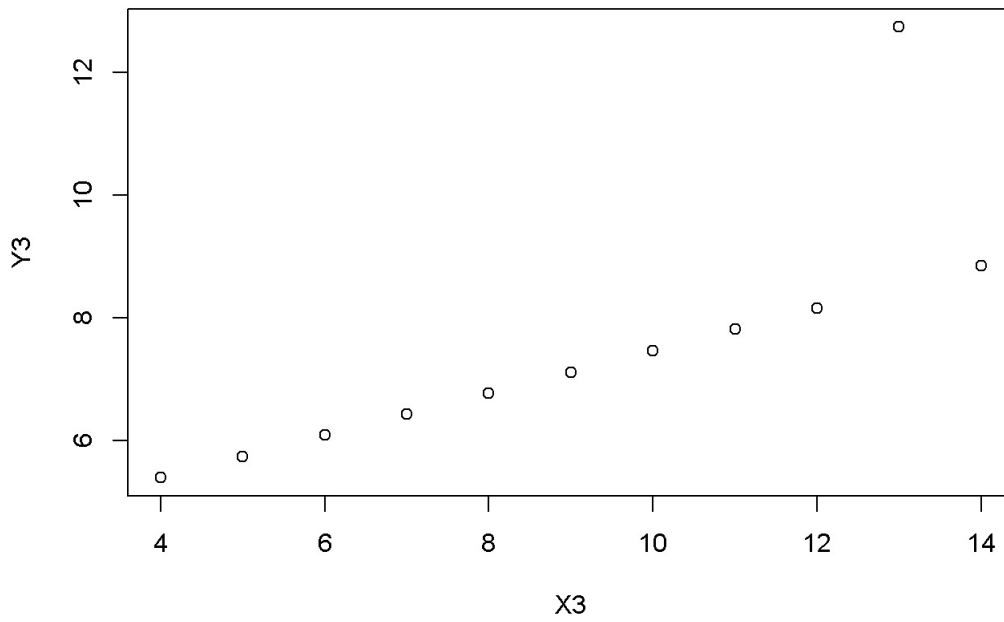
```
#Exercise 3
plot(ds1)
```



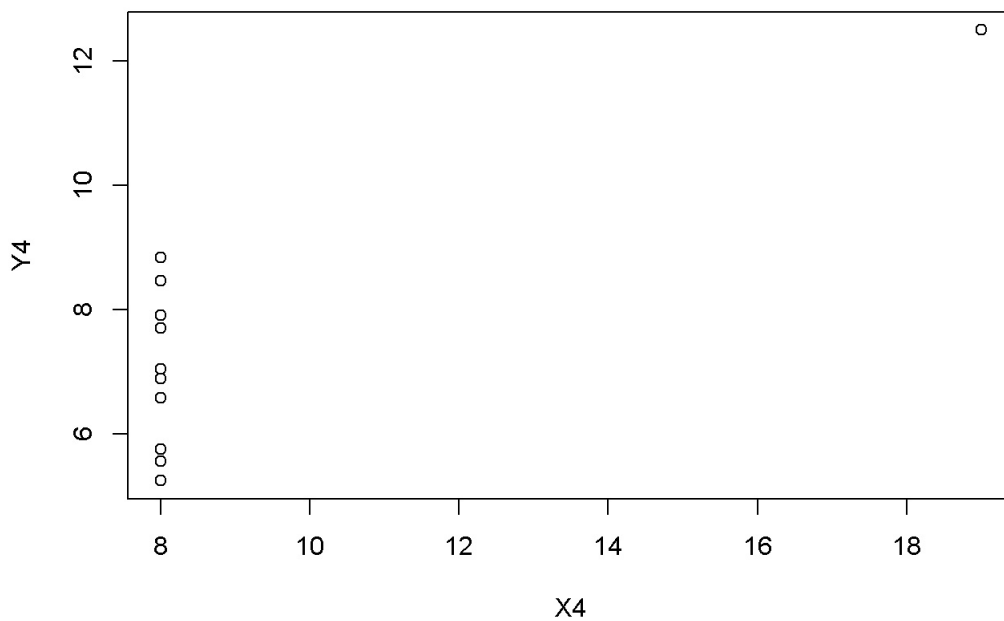
```
plot(ds2)
```



```
plot(ds3)
```



```
plot(ds4)
```



```
#Exercise 4
#Mean of X in all plots is same = 9
#Mean of Y in all plots is same = 7.5 (upto 1 decimal place)
#Plot for first data set is a linear regression or relationship
#Plot for second data set is a positive relationship but with 3 outliers towards the end which have a negative
relationship so it is a non-linear relationship
#Plot for third data set is a positive relationship with 1 outlier
#Plot for fourth data set is a null relationship with 1 outlier

#Exercise 5
#Importance of Exploratory Analysis - Looking at a set of data graphically before starting to analyze according
to a particular type of relationship is important.
#
# If a dataset has equal mean on both axis does not mean it will have a lin
ear relationship but still be treated as a linear model based on mathematical calculations as both X and Y have
same mean.
```