1.Covariance and correlation

Children of three ages are asked to indicate their preference for three photographs of adults. Do the data suggest that there is a significant relationship between age and photograph preference? What is wrong with this study?

	Photograph:		
Age of child	A	B	C
5-6 years:	18	22	20
7-8 years:	2	28	40
9-10 years:	20	10	40

1. Use cov() to calculate the sample covariance between B and C.
2. Use another call to cov() to calculate the sample covariance matrix for the preferences.
3. Use cor() to calculate the sample correlation between B and C.
4. Use another call to cor() to calculate the sample correlation matrix for the preferences.

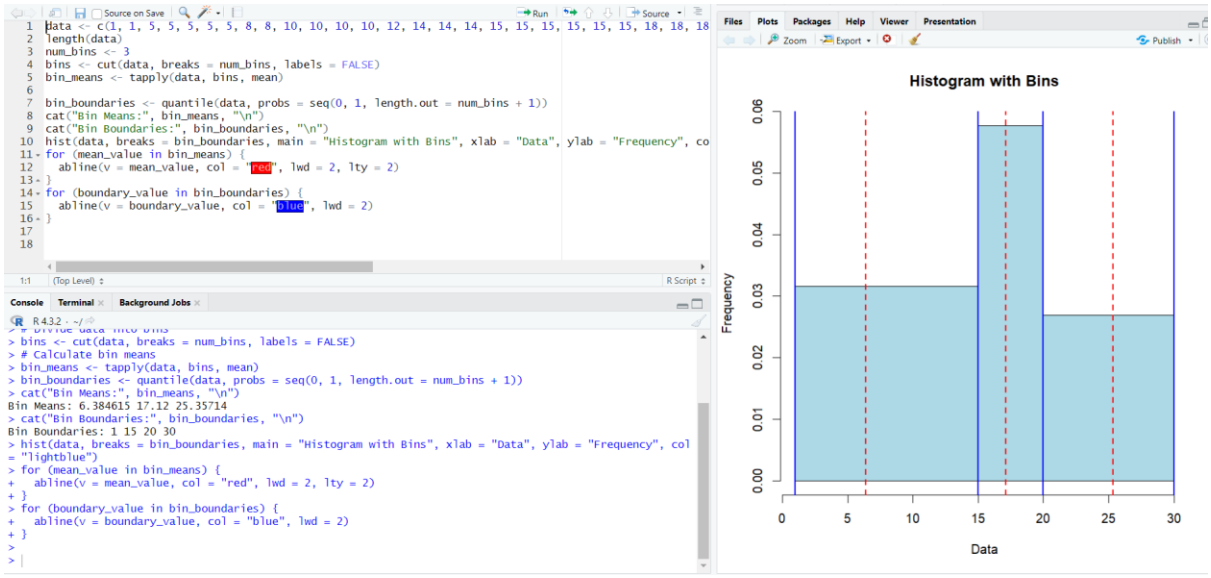
```
b <-c(22,28,10)
c <-c(20,40,40)
s1 <- cov(b,c)
s2 <-cov((cbind(b,c)))
s3 <-cor(b,c)
s4 <-cor((cbind(b,c)))
print(s1)
print(s2)
print(s3)
print(s4)
```

```
> s1 <- cov(b,c)
> s2 <-cov((cbind(b,c)))
> s3 <-cor(b,c)
> s4 <-cor((cbind(b,c)))
> print(s1)
[1] -20
> print(s2)
    b        c
b  84 -20.0000
c -20 133.3333
> print(s3)
[1] -0.1889822
> print(s4)
           b          c
b  1.0000000 -0.1889822
c -0.1889822  1.0000000
>
```

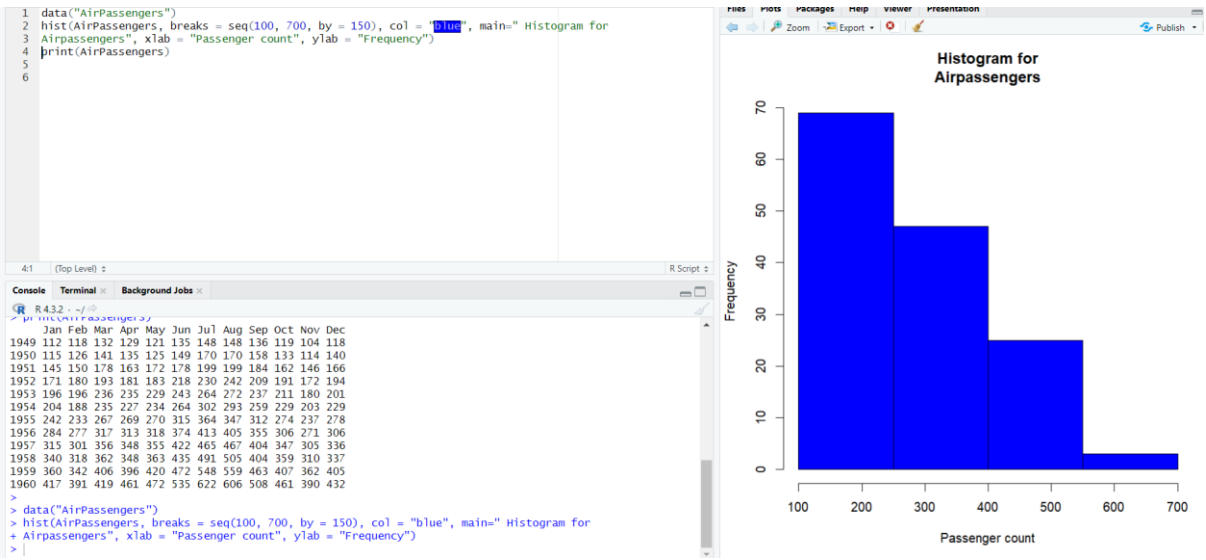
2. Imagine that you have selected data from the All Electronics data warehouse for analysis. The data set will be huge! The following data are a list of All Electronics prices for commonly sold items (rounded to the nearest dollar). The numbers have been sorted: 1, 1, 5, 5, 5, 5, 5, 8, 8, 10, 10, 10, 10, 12, 14, 14, 14, 15, 15, 15, 15, 15, 15, 18, 18, 18, 18, 18, 18, 18, 18, 20, 20, 20, 20, 20, 20, 20, 21, 21, 21, 21, 25, 25, 25, 25, 25, 28, 28, 30,

30, 30.

(i) Partition the dataset using an equal-frequency partitioning method with bin equal to 3
(ii) apply data smoothing using bin means and bin boundary.
(iii) Plot Histogram for the above frequency division



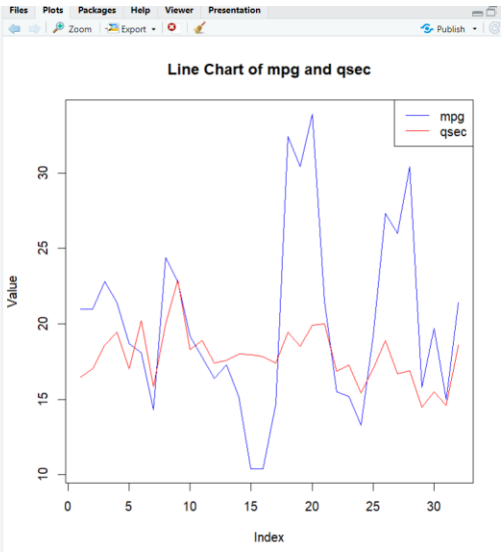
3. Make a histogram for the “AirPassengers “dataset, start at 100 on the x-axis, and from values 200 to 700, make the bins 150 wide



4. Obtain Multiple Lines in Line Chart using a single Plot Function in R.Use attributes“mpg”and“qsec”of the dataset “mtcars”

```r
# Load the dataset
data(mtcars)

# Create an empty plot with the first set of data
plot(mtcars$mpg, type = "l", col = "blue", ylim = range(c(mtcars$mpg, mtcars$qsec)),
     xlab = "Index", ylab = "Value")

# Add the second set of data as another line
lines(mtcars$qsec, col = "red")

# Add a legend
legend("topright", legend = c("mpg", "qsec"), col = c("blue", "red"), lty = 1)

# Add a title
title(main = "Line Chart of mpg and qsec
```

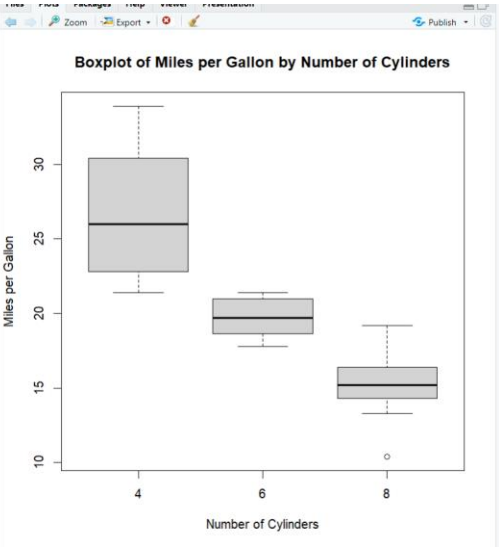
```
> # Load the dataset
> data(mtcars)
> # Create an empty plot with the first set of data
> plot(mtcars$mpg, type = "l", col = "blue", ylim = range(c(mtcars$mpg, mtcars$qsec)),
+      xlab = "Index", ylab = "Value")
> # Add the second set of data as another line
> lines(mtcars$qsec, col = "red")
> # Add a legend
> legend("topright", legend = c("mpg", "qsec"), col = c("blue", "red"), lty = 1)
> # Add a title
> title(main = "Line Chart of mpg and qsec")
```



5. **Create a Boxplot graph for the relation between "mpg"(miles per galloon) and "cyl"(number of Cylinders) for the dataset "mtcars" available in R Environment.**

```r
# Load the dataset
data(mtcars)

# Create a boxplot
boxplot(mpg ~ cyl, data = mtcars,
        xlab = "Number of Cylinders",
        ylab = "Miles per Gallon",
        main = "Boxplot of Miles per Gallon by Number of Cylinders")
```

```
> # Load the dataset
> data(mtcars)
> # Create a boxplot
> boxplot(mpg ~ cyl, data = mtcars,
+         xlab = "Number of Cylinders",
+         ylab = "Miles per Gallon",
+         main = "Boxplot of Miles per Gallon by Number of Cylinders")
```



6. Assume the Tennis coach wants to determine if any of his team players are scoring outliers. To visualize the distribution of points scored by his players, then how can he decide to develop the box plot? Give suitable example using Boxplot visualization technique.

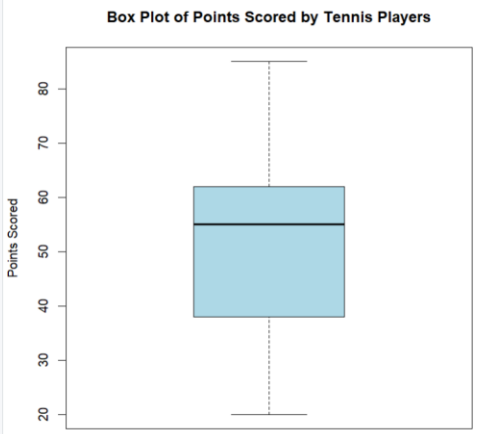
```r
score <- c(20, 25, 30, 32, 35, 38, 40, 45, 50, 52, 55, 56, 58, 59, 60, 62, 65, 70, 75, 80, 85)
boxplot(score, col = "lightblue", main = "Box Plot of Points Scored by Tennis Players", ylab =
          "Points Scored")
```

4:1 (Top Level) R Script

Console Terminal Background Jobs

R 4.3.2 · ~/

```
> score <- c(20, 25, 30, 32, 35, 38, 40, 45, 50, 52, 55, 56, 58, 59, 60, 62, 65, 70, 75, 80, 85)
> boxplot(score, col = "lightblue", main = "Box Plot of Points Scored by Tennis Players", ylab =
+           "Points Scored")
>
>
```

Files Plots Packages Help Viewer Presentation

Zoom Export Publish