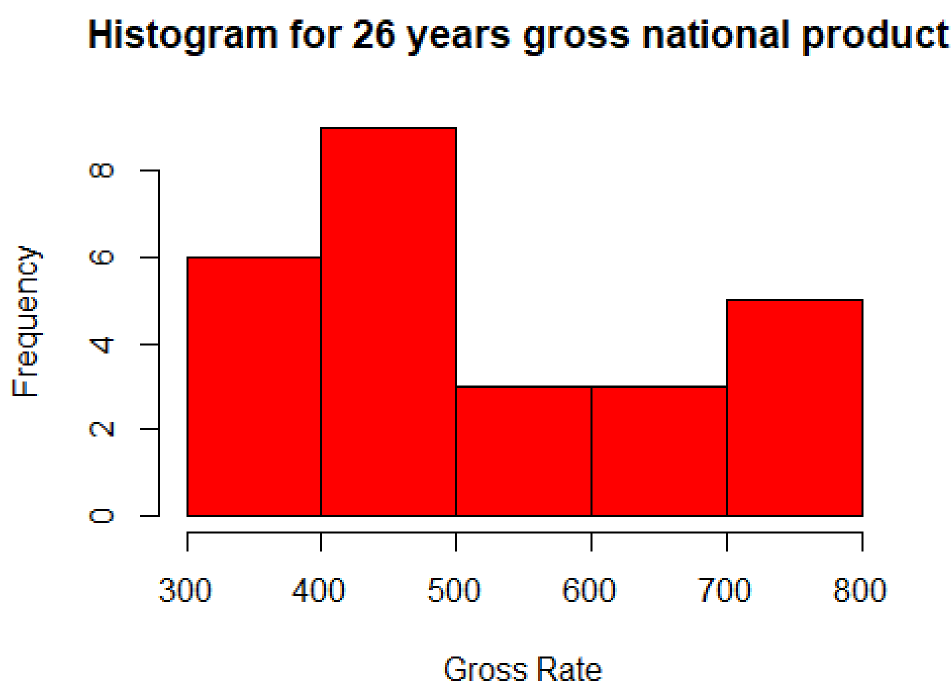
```r
# Q1 File gross.dat contains gross national product y in real dollars for 26 recent years,
# x = 1,2,...,26. Use R expressions to do the followings

# a. Read data into R.
gross = read.table("gross.dat", header = TRUE)
grossProduct = gross$y;
grossProduct
```

```
##  [1] 309.9 323.7 324.1 355.3 383.4 395.1 412.8 407.0 438.0 446.1 452.5
## [12] 447.3 475.9 487.7 497.2 529.8 551.0 581.1 617.8 658.1 675.2 706.6
## [23] 725.6 722.5 745.4 790.7
```

```r
# b. Plot the histogram of y.
hist(grossProduct, col="red", xlab="Gross Rate", main="Histogram for 26 years
 gross national product ")
```




```r
# c. Calculate the sample mean, the sample standard deviation1 and median of y.
#   i. Mean of gross product
meanGP = mean(grossProduct)
meanGP
```

```
## [1] 517.6846
```

```r
#   ii. Standard Deviation of Gross Product
varGP = sum((grossProduct - meanGP)^2)/(length(grossProduct) - 1)
sdGP = sqrt(varGP)
sdGP
```

```
## [1] 145.8477
```

```r
#   iii. median of Gross Product
medianGP = median(grossProduct)
medianGP
```

```
## [1] 481.8
```

```r
# d.  The coding of a data set refers to the operation of subtracting (or add
ing)
#     a constant to each observation and then dividing (multiplying) by anoth
er constant.
#     The coding by subtracting the sample mean of data and dividing by the s
ample
#     standard deviation is called standardization and the coded data is call
ed standardized data.
#     Standardize y and calculate the mean and the variance of the standardiz
ed y

## Standardize y
standardizedGP = (grossProduct - meanGP)/sdGP
standardizedGP
```

```
##  [1] -1.42466833 -1.33004909 -1.32730650 -1.11338473 -0.92071801
##  [6] -0.84049735 -0.71913789 -0.75890540 -0.54635492 -0.49081754
## [11] -0.44693615 -0.48258978 -0.28649483 -0.20558852 -0.14045209
## [16]  0.08306873  0.22842583  0.43480548  0.68643782  0.96275343
## [21]  1.07999901  1.29529207  1.42556494  1.40430990  1.56132299
## [26]  1.87192094
```

```r
## Mean of Standardize y
meanStandardizedGP = mean(standardizedGP)
meanStandardizedGP
```

```
## [1] -2.910416e-16
```

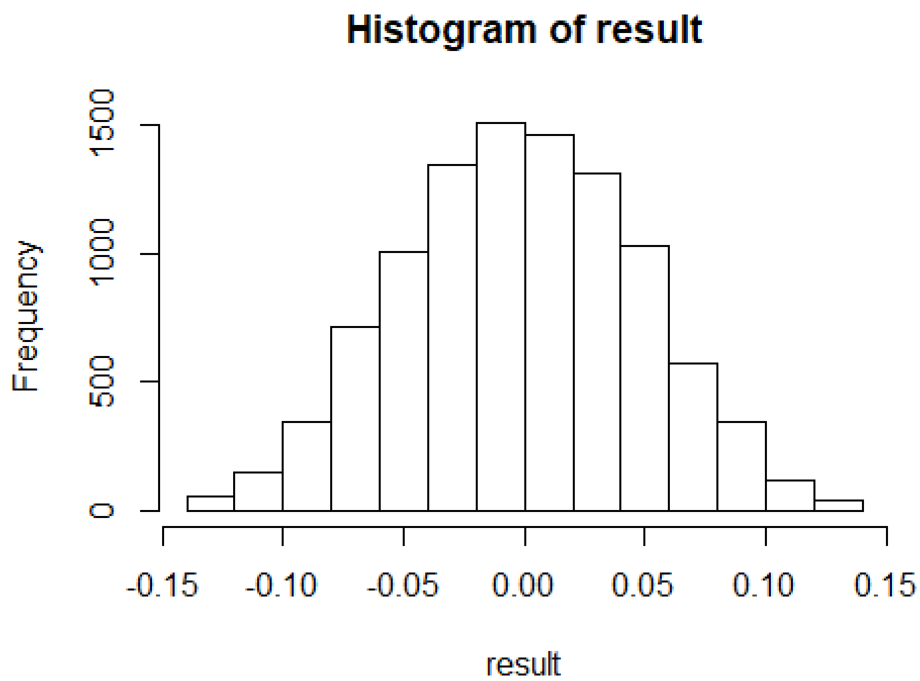
```r
## Variance of Standardize y
varianceStandardizedGP = var(standardizedGP)
varianceStandardizedGP
```

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

```
# Q2 J.J. Thomson (1856-1940), discovered the electron while investigating th
e basic nature of cathode rays.
# In laboratory experiments Thomson isolated negatively charged particles for
 which he could determine
# the mass-charge ratio. This ratio appeared to be constant over a wide varie
ty of experimental conditions
# and to be a characteristic of these new particles. Thomson obtained the fol
lowing results with two
# different cathode ray tubes, using air as the gas:

# Tube 1 | 0.57 0.34 0.43 0.32 0.48 0.40 0.40
# Tube 2 | 0.53 0.47 0.47 0.51 0.63 0.61 0.48

# (a) Do the two tubes appear to produce consistent results?
tube1 = c(.57, .34, .43, .32, .48, .40, .40)
tube2 = c(.53, .47, .47, .51, .63, .61, .48)

meanTube1 = mean(tube1)
meanTube2 = mean(tube2)

resultDiff = meanTube2 - meanTube1

## Result of 2 tubes mean
resultDiff
```

```
## [1] 0.1085714
```

```
allTogether = c(tube1, tube2)

result=0
for (i in 1:10000) {
  mixedTogether = sample(allTogether)
  tube1R = mixedTogether[1:7]
  tube2R = mixedTogether[8:14]
  result[i]= mean(tube1R) - mean(tube2R)
}

hist(result)
```




```
## After running above calculation, we came to conclusion that Results are NO
T CONSISTENT,
## Since the difference of both tube resulted 0.10857, while populating histo
gram for 10000 values we found that
## this is far too right in histogram.
```

```
# Q3 An orbiting satellite has 3 panels of solar cells, all of which must be
active to provide an adequate
# power output. The panels function independently of one another. The chance
that a single panel will
# fail during the mission is 0.02. What is the probability that there will be
 adequate power output during
# the entire mission time? (This probability is called the reliability of the
 system.)

# success probability out of failure for one cell
successProb = 1 - .02

# if any one fails, it will not have adequate power, So the answer is
dbinom(3, 3, successProb)
```

```
## [1] 0.941192
```

```r
# Q4 In a study of the coexistence of two types of insects, let X and Y denot
e the number of type A and type
# B insects, respectively, that reside on the same plant. From observations o
f a large number of plants,
# suppose that the following joint probability distribution is assessed for t
he insect counts per plant.


# x→y↓| 1    2    3    4

#  0  | 0    0.05 0.05 0.10
#  1  | 0.08 0.15 0.10 0.10
#  2  | 0.20 0.12 0.05 0

# marginal X | .28 .32 .20 .20
# marginal Y | .20 .43 .37


# (a) Find the probability that there are more type B insects than type A ins
ects on a plant.

# Omitting (0,1) (2,4) Plants, since they both are 0 for Type B and Type A in
sects

# P(Y|X) for (1, 1) = X intersection Y / X
.08/.28 > .28

## [1] TRUE

# P(Y|X) for (2, 1) = X intersection Y / X
.20/.28 > .28

## [1] TRUE

# P(Y|X) for (0, 2) = X intersection Y / X
.05/.32 > .32

## [1] FALSE

# P(Y|X) for (1, 2) = X intersection Y / X
.15/.32 > .32

## [1] TRUE

# P(Y|X) for (2, 2) = X intersection Y / X
.12/.32 > .32

## [1] TRUE

# P(Y|X) for (0, 3) = X intersection Y / X
.05/.20 > .20
```

```
## [1] TRUE

# P(Y|X) for (1, 3) = X intersection Y / X
.10/.20 > .20

## [1] TRUE

# P(Y|X) for (2, 3) = X intersection Y / X
.05/.20 > .20

## [1] TRUE

# P(Y|X) for (2, 2) = X intersection Y / X
.12/.32 > .32

## [1] TRUE

# P(Y|X) for (0, 4) = X intersection Y / X
.10/.20 > .20

## [1] TRUE

# P(Y|X) for (1, 4) = X intersection Y / X
.10/.20 > .20

## [1] TRUE

## Since only a plant (0,2) has more Type A insects Probability then Type B
insects indicated by False in results, While Rest of the plants have Type B
insects Probability indicated by TRUE in results

# (b) Compute μx, μy, σx, σy and Cov(X, Y).

xi = 1:4
x = c(.28, .32, .20, .20)
yj = 0:2
y = c(.20, .43, .37)

# μx
mx = sum(xi*x)
mx

## [1] 2.32

# μy
my = sum(yj*y)
my

## [1] 1.17
```

```
# σx
varx = sum((xi^2)*x) - mx^2
varx

## [1] 1.1776
```

```
# σy
vary = sum((yj^2)*y) - my^2
vary

## [1] 0.5411
```

```
mxy = sum(
  1*1*.08, 1*2*.20,
  2*1*.15, 2*2*.12,
  3*1*.10, 3*2*.05,
  4*1*.10, 4*2*0
)

# Cov(X, Y)
covXY = mxy - mx*my
covXY

## [1] -0.4544
```

```
# Q5 The medical records of the male diabetic patents reporting to a clinic d
uring one year provide the
# following percentage:
#            | Light case            | Serious case

# Age of     | Diabetes in parents | Diabetes in parents
# patient    | Yes       |        No | Yes       |        No

# Below 40   | 15                 10 | 8         |         2

# Above 40   | 15                 20 | 20        |        10

# Suppose a patient is chosen at random from this group, and the events A, B
and C are defined
# A: He has a serious case.
# B: He is below 40.
# C: His parents are diabetic

## Total number of patients are 100, so every element divided by 100 to get P
robability
```

```
# (a) Find the probabilities P(A), P(B), P(A ∩ B), P(A ∩ B ∩ C).

## P(A)
.4

## [1] 0.4


## P(B)
.35

## [1] 0.35


## P(A ∩ B)
.1

## [1] 0.1


## P(A ∩ B ∩ C)
.08

## [1] 0.08


# (b) Describe the following events verbally and find their probabilities:
## (i) A' ∩ B'
## patients that are not serious and not below 40
.35

## [1] 0.35


## (ii) A'∪C'
## Patient that are not serious or with parents that are not diabetic
.72

## [1] 0.72


## (iii) A' ∩ B ∩ C'
## patients that are not serious, below 40 and parents are not diabetic
.1

## [1] 0.1
```