Ketaki Mahajan

A – 2 / 16014022050

Tutorial 1: Correlation & Regression (19 / 01 / 2024)

Q1.

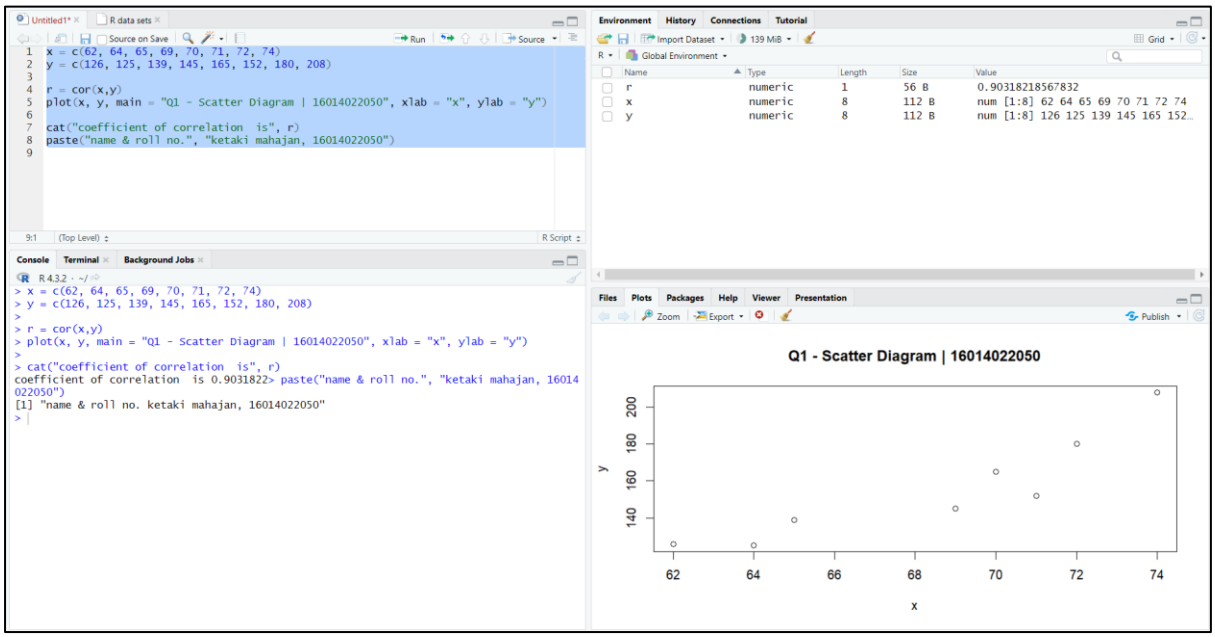
Draw scatter diagram and determine the coefficient of correlation for the following data:

X	62	64	65	69	70	71	72	74
Y	126	125	139	145	165	152	180	208

Code –

```
x = c(62, 64, 65, 69, 70, 71, 72, 74)
y = c(126, 125, 139, 145, 165, 152, 180, 208)

r = cor(x,y)
plot(x, y, main = " Q1 - Scatter Diagram | 16014022050", xlab = "x", ylab = "y")

cat("coefficient of correlation  is", r)
paste("name & roll no.", "ketaki mahajan, 16014022050")
```

Output –



```
> x = c(62, 64, 65, 69, 70, 71, 72, 74)
> y = c(126, 125, 139, 145, 165, 152, 180, 208)
>
> r = cor(x,y)
> plot(x, y, main = "Q1 - Scatter Diagram | 16014022050", xlab = "x", ylab = "y")
>
> cat("coefficient of correlation  is", r)
coefficient of correlation  is 0.9031822> paste("name & roll no.", "ketaki mahajan, 16014
022050")
[1] "name & roll no. ketaki mahajan, 16014022050"
>
```

Q2.

Obtain the equation of the line of y on x.

Estimate y when x is 73.

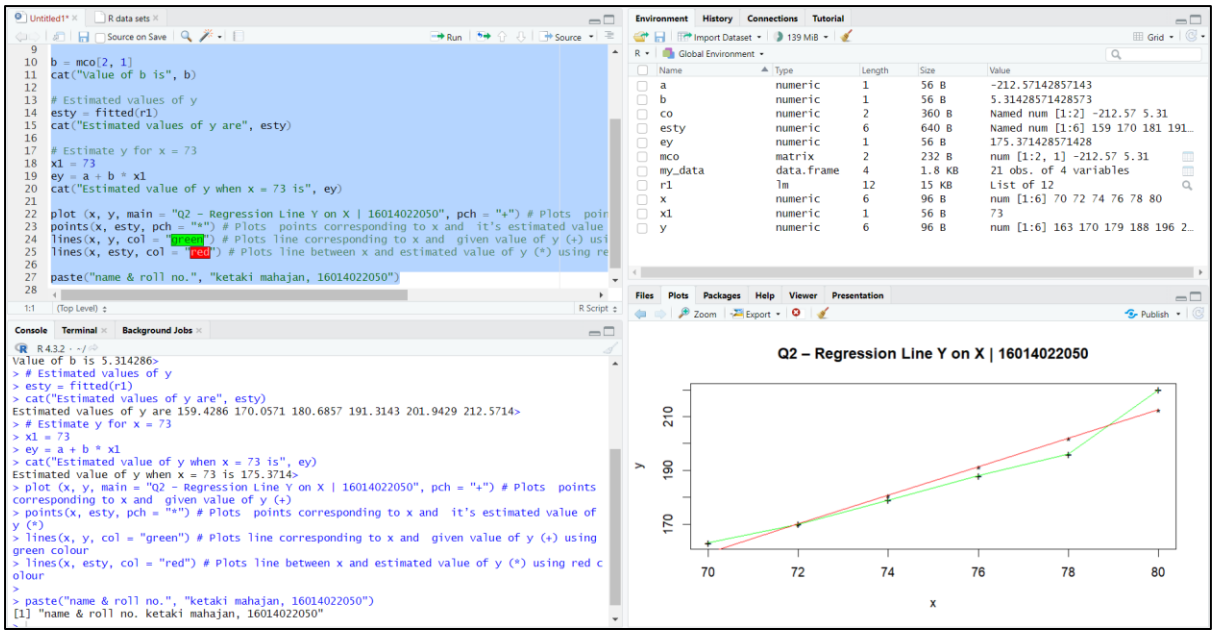
Plot equation of regression line of y on x.

X	70	72	74	76	78	80
Y	163	170	179	188	196	220

Code –

```
x = c(70, 72, 74, 76, 78, 80)
y = c(163, 170, 179, 188, 196, 220)

r1 = lm(y ~ x)  # Regression of y on x
co = coef(r1)  # Coefficients (a, b) in the equation y = a + b*x
mco = matrix(co)  # Column matrix of coefficients (a, b)
a = mco[1, 1]
cat("Constant term a is", a)

b = mco[2, 1]
cat("Value of b is", b)

# Estimated values of y
esty = fitted(r1)
cat("Estimated values of y are", esty)

# Estimate y for x = 73
x1 = 73
ey = a + b * x1
cat("Estimated value of y when x = 73 is", ey)

plot (x, y, main = "Q2 – Regression Line Y on X | 16014022050", pch = "+") # Plots points corresponding to x and  given value of y (+)
points(x, esty, pch = "*") # Plots  points corresponding to x and  it's estimated value of y (*)
lines(x, y, col = "green") # Plots line corresponding to x and  given value of y (+) using green colour
lines(x, esty, col = "red") # Plots line between x and estimated value of y (*) using red colour

paste("name & roll no.", "ketaki mahajan, 16014022050")
```

Output –




```
> x = c(70, 72, 74, 76, 78, 80)
> y = c(163, 170, 179, 188, 196, 220)
>
> r1 = lm(y ~ x)  # Regression of y on x
> co = coef(r1)  # Coefficients (a, b) in the equation y = a + b*x
> mco = matrix(co)  # Column matrix of coefficients (a, b)
> a = mco[1, 1]
> cat("Constant term a is", a)
Constant term a is -212.5714>
> b = mco[2, 1]
> cat("Value of b is", b)
Value of b is 5.314286>
> # Estimated values of y
> esty = fitted(r1)
> cat("Estimated values of y are", esty)
Estimated values of y are 159.4286 170.0571 180.6857 191.3143 201.9429 212.571
4>
> # Estimate y for x = 73
> x1 = 73
> ey = a + b * x1
> cat("Estimated value of y when x = 73 is", ey)
Estimated value of y when x = 73 is 175.3714>
> plot (x, y, main = "Q2 – Regression Line Y on X | 16014022050", pch = "+") #
Plots  points corresponding to x and  given value of y (+)
> points(x, esty, pch = "*") # Plots  points corresponding to x and  it's esti
mated value of y (*)
> lines(x, y, col = "green") # Plots line corresponding to x and  given value
of y (+) using green colour
> lines(x, esty, col = "red") # Plots line between x and estimated value of y
(*) using red colour
>
> paste("name & roll no.", "ketaki mahajan, 16014022050")
[1] "name & roll no. ketaki mahajan, 16014022050"
>
```

Q3.

Find the equations of lines of regression of x on y for the following data.

Estimate x when y is 70.

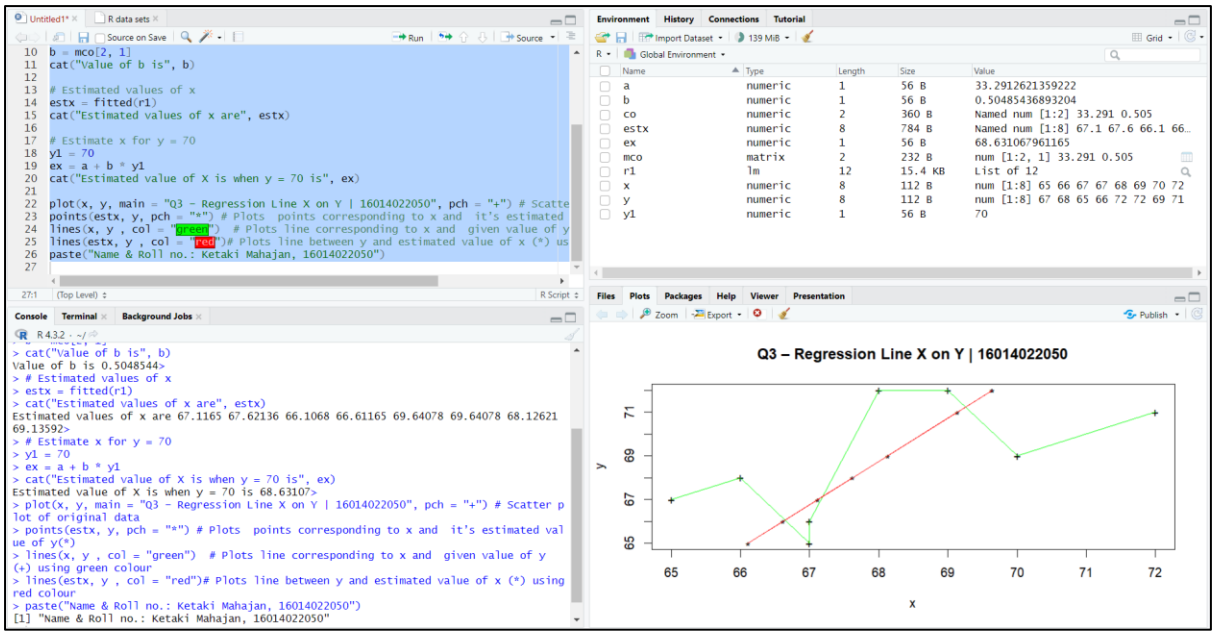
Plot line of regression of x on y.

X	65	66	67	67	68	69	70	72
Y	67	68	65	66	72	72	69	71

Code –

```
x = c(65, 66, 67, 67, 68, 69, 70, 72)
y = c(67, 68, 65, 66, 72, 72, 69, 71)

r1 = lm(x ~ y)  # Regression of x on y
co = coef(r1)  # Coefficients (a, b) in the equation x = a + b*y
mco = matrix(co)  # Column matrix of coefficients (a, b)
a = mco[1, 1]
cat("Constant term a is", a)

b = mco[2, 1]
cat("Value of b is", b)

# Estimated values of x
estx = fitted(r1)
cat("Estimated values of x are", estx)

# Estimate x for y = 70
y1 = 70
ex = a + b * y1
cat("Estimated value of X is when y = 70 is", ex)

plot(x, y, main = “Q3 – Regression Line X on Y | 16014022050”, pch = "+") # Scatter plot of original data
points(estx, y, pch = "*") # Plots  points corresponding to x and  it’s estimated value of y(*)
lines(x, y , col = "green")  # Plots line corresponding to x and  given value of y   (+) using green colour
lines(estx, y , col = "red")# Plots line between y and estimated value of x (*) using red colour
paste("Name & Roll no.: Ketaki Mahajan, 16014022050")
```

Output –




```
> x = c(65, 66, 67, 67, 68, 69, 70, 72)
> y = c(67, 68, 65, 66, 72, 72, 69, 71)
>
> r1 = lm(x ~ y)  # Regression of x on y
> co = coef(r1)  # Coefficients (a, b) in the equation x = a + b*y
> mco = matrix(co)  # Column matrix of coefficients (a, b)
> a = mco[1, 1]
> cat("Constant term a is", a)
Constant term a is 33.29126>
> b = mco[2, 1]
> cat("Value of b is", b)
Value of b is 0.5048544>
> # Estimated values of x
> estx = fitted(r1)
> cat("Estimated values of x are", estx)
Estimated values of x are 67.1165 67.62136 66.1068 66.61165 69.64078 69.64078 68.12621 6
9.13592>
> # Estimate x for y = 70
> y1 = 70
> ex = a + b * y1
> cat("Estimated value of X is when y = 70 is", ex)
Estimated value of X is when y = 70 is 68.63107>
> plot(x, y, main = "Q3 - Regression Line X on Y | 16014022050", pch = "+") # Scatter plo
t of original data
> points(estx, y, pch = "*") # Plots  points corresponding to x and  it's estimated value
of y(*)
> lines(x, y , col = "green")  # Plots line corresponding to x and  given value of y
(+) using green colour
> lines(estx, y , col = "red")# Plots line between y and estimated value of x (*) using r
ed colour
> paste("Name & Roll no.: Ketaki Mahajan, 16014022050")
[1] "Name & Roll no.: Ketaki Mahajan, 16014022050"
>
```