Question 4 f

R Script

```r
x <- c(10, 10, 13, 13, 18, 19, 22)
y <- c(66, 66, 108, 106, 161, 166, 88)

# QUESTION 4 - A
cat("Question 4 - A\n")
print_line()

# a) Find and interpret the linear correlation coefficient.
print(summary(lm(y ~ x)))
cat("linear correlation coefficient:", cor(x, y), "\n")
cat("Correlation test:\n")
print(cor.test(x, y))

print_line()

# QUESTION 4 - B
cat("Question 4 - B\n")
print_line()

# b) Find and interpret the coefficient of determination.
cat("Coefficient of determination:", summary(lm(y ~ x))$r.squared, "\n")
print_line()

# c) Find the least-squares estimates for the regression line
cat("Question 4 - C\n")
print_line()
cat("Least-squares estimates for the regression line:\n")
coefficients <- summary(lm(y ~ x))$coefficients
print(coefficients)

# Interpret the slope and intercept
slope <- coefficients[2, 1]
intercept <- coefficients[1, 1]
cat("Slope: y increases by", slope, "per unit of x.\n")
cat("Intercept: y is", intercept, "when x = 0.\n")
print_line()

# d) Predict y if x = 15 and compare to y = 150
cat("Question 4 - D\n")
```

```r
print_line()
x_pred <- 15
y_pred <- intercept + slope * x_pred
cat("Predicted y for x = 15:", y_pred, "\n")
actual_y <- 150
residual <- actual_y - y_pred
cat("Residual (actual - predicted):", residual, "\n")
print_line()

# e) Construct a 90% confidence interval for the slope and interpret
cat("Question 4 - E\n")
print_line()
ci_slope <- confint(lm(y ~ x), level = 0.9)["x", ]
cat("90% Confidence Interval for the slope:",
  ci_slope[1], "to", ci_slope[2], "\n"
)
cat("We are 90% confident that the true increase
in y per unit increase in x lies within this range.\n")

# Compare to hypothesis testing
p_value <- coefficients[2, 4]
cat("P-value for the slope:", p_value, "\n")
# if p val is less than 0.1, we reject the null
#hypothesis that the slope is equal to 0
print_line()


#                    5 %      95 %
# (Intercept) -67.871301 131.89078
# x            -1.282465  11.50974
```

Output

```
Question 4 - A
---------------------------------------

Call:
lm(formula = y ~ x)

Residuals:
      1       2       3       4       5       6       7
-17.146 -17.146   9.513   7.513  36.945  36.831 -56.510

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   32.010     49.568   0.646    0.547
x              5.114      3.174   1.611    0.168

Residual standard error: 36.47 on 5 degrees of freedom
Multiple R-squared:  0.3417,        Adjusted R-squared:   0.21
F-statistic: 2.595 on 1 and 5 DF,  p-value: 0.1681

linear correlation coefficient: 0.584555
Correlation test:

        Pearson's product-moment correlation

data:  x and y
t = 1.611, df = 5, p-value = 0.1681
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.3010079  0.9287665
sample estimates:
     cor
0.584555

---------------------------------------
Question 4 - B
---------------------------------------
Coefficient of determination: 0.3417046
---------------------------------------
Question 4 - C
```

```
----------------------------------------
Least-squares estimates for the regression line:
             Estimate Std. Error  t value  Pr(>|t|)
(Intercept) 32.009740  49.567565 0.645780 0.5468855
x            5.113636   3.174168 1.611017 0.1680924
Slope: y increases by 5.113636 per unit of x.
Intercept: y is 32.00974 when x = 0.
----------------------------------------
Question 4 - D
----------------------------------------
Predicted y for x = 15: 108.7143
Residual (actual - predicted): 41.28571
----------------------------------------
Question 4 - E
----------------------------------------
90% Confidence Interval for the slope: -1.282465 to 11.50974
We are 90% confident that the true increase
in y per unit increase in x lies within this range.
P-value for the slope: 0.1680924
----------------------------------------
```