The weight and systolic blood pressure of 26 randomly selected males in the age group 25-30 are shown below. Assume that weight and blood pressure are jointly normally distributed.

a. Find a regression line relating systolic Blood pressure to Weight

b. Estimate the correlation coefficient

c. Test the hypothesis that correlation coefficient (rho) = 0.

d. Find a 95% confidence interval for correlation coefficient.

```
# Run the file to load the dataframe

df = data.frame(subject=c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,
21,22,23,24,25,26),
                weight=c(165,167,180,155,212,175,190,210,200,149,158,169,170,
172,159,168,174,183,215,195,180,143,240,235,192,187),
                systolic_BP=c(130,133,150,128,151,146,150,140,148,125,133,135
,150,153,128,132,149,158,150,163,156,124,170,165,160,159))
```

A) Find a regression line relating systolic Blood pressure to Weight

```
# To visualize the dataframe as a Table
View(df)

model1 <- lm(systolic_BP~weight,data=df)
model1

##
## Call:
## lm(formula = systolic_BP ~ weight, data = df)
##
## Coefficients:
## (Intercept)       weight
##     69.1044       0.4194
```

Hence, the regression equation of systolic BP on weight is: Systolic_BP = 69.1044 + 0.4194 * weight

B) Estimate the correlation coefficient

```
cor(df$weight, df$systolic_BP)

## [1] 0.7734903
```

We can observe that Weight and Systolic BP are highly correlated.

C) Test the hypothesis that correlation coefficient (rho) = 0.

```
model1 <- lm(systolic_BP~weight,data=df)
summary(model1)
```

```
##
## Call:
## lm(formula = systolic_BP ~ weight, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -17.182  -6.485  -2.519   8.926  12.143
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 69.10437   12.91013   5.353 1.71e-05 ***
## weight       0.41942    0.07015   5.979 3.59e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.681 on 24 degrees of freedom
## Multiple R-squared:  0.5983, Adjusted R-squared:  0.5815
## F-statistic: 35.74 on 1 and 24 DF,  p-value: 3.591e-06
```

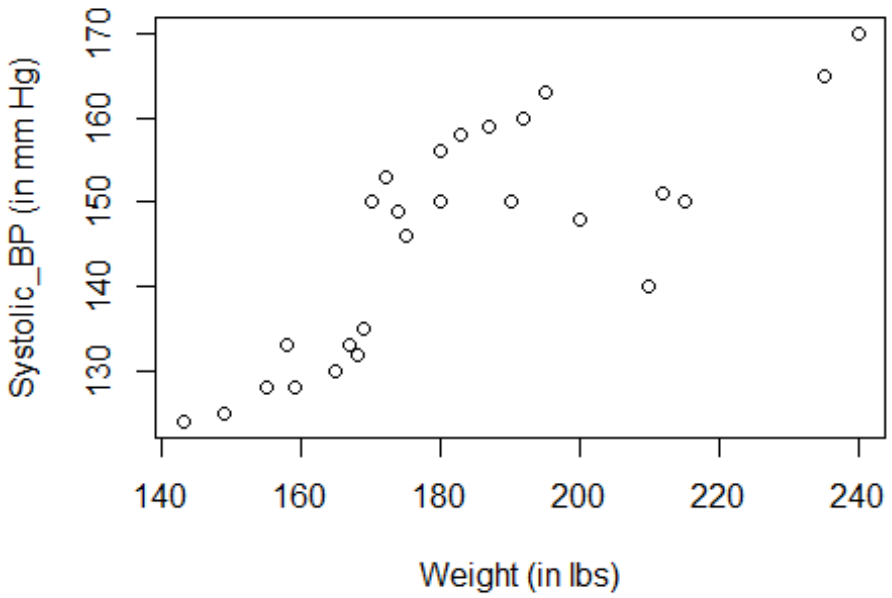
```
fitted(model1)
```

```
##        1        2        3        4        5        6        7        8
## 138.3079 139.1467 144.5991 134.1137 158.0204 142.5020 148.7933 157.1816
##        9       10       11       12       13       14       15       16
## 152.9874 131.5972 135.3720 139.9855 140.4050 141.2438 135.7914 139.5661
##       17       18       19       20       21       22       23       24
## 142.0826 145.8574 159.2786 150.8903 144.5991 129.0807 169.7640 167.6669
##       25       26
## 149.6321 147.5350
```

```
residuals(model1)
```

```
##           1           2           3           4           5           6
##  -8.3078813  -6.1467117   5.4008907  -6.1137292  -7.0203958   3.4979667
##           7           8           9          10          11          12
##   1.2067387 -17.1815654  -4.9874134  -6.5972380  -2.3719749  -4.9855421
##          13          14          15          16          17          18
##   9.5950427  11.7562123  -7.7913901  -7.5661269   6.9173819  12.1426451
##          19          20          21          22          23          24
##  -9.2786414  12.1096626  11.4008907  -5.0807468   0.2359785  -2.6669455
##          25          26
##  10.3679082  11.4649843
```

```
plot(df$weight,df$systolic_BP,
     xlab = "Weight (in lbs)",
     ylab = "Systolic_BP (in mm Hg)")
```




```
c = cor.test(df$weight,df$systolic_BP)
c

##
##  Pearson's product-moment correlation
##
## data:  df$weight and df$systolic_BP
## t = 5.9786, df = 24, p-value = 3.591e-06
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.5513214 0.8932215
## sample estimates:
##       cor
## 0.7734903
```

Default confidence interval: 95%

D) Find a 95% confidence interval for correlation coefficient.

```
d = cor.test(df$weight,df$systolic_BP,level=0.95)
d

##
##  Pearson's product-moment correlation
##
## data:  df$weight and df$systolic_BP
## t = 5.9786, df = 24, p-value = 3.591e-06
```

```
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.5513214 0.8932215
## sample estimates:
##       cor
## 0.7734903
```