LINEAR REGRESSION ON CHOLESTEROL DATA

Using R studio

```
> library(haven)
> fat <- read.csv("D:\reachout analytics\case studies(roba)\Case Studies 74\Bodyfat\Cholestrol.csv")
Error: '\c' is an unrecognized escape in character string starting ""D:\reachout analytics\c"
> View(fat)
Error in View : object 'fat' not found
> library(haven)
> fat <- read.csv("D:\reachout analytics\case studies(roba)\Case Studies 74\Bodyfat\Cholestrol.xlsx")
Error: '\c' is an unrecognized escape in character string starting ""D:\reachout analytics\c"
> View(fat)
Error in View : object 'fat' not found
> library(readxl)
> Cholestrol <- read_excel("D:/reachout analytics/case studies(roba)/Case Studies 74/Bodyfat/Cholestrol.xlsx")
> View(Cholestrol)
> summary(cholestrol)
Error in summary(cholestrol) : object 'cholestrol' not found
> summary(Cholestrol)
    Body_FAT           Age            weight           Height         Neck_Circ        Chest_Circ
 Min.   :-1.30    Min.   :20.00   Min.   : 54.80   Min.   :160.6   Min.   :29.1    Min.   : 83.4
 1st Qu.:14.80    1st Qu.:38.00   1st Qu.: 74.10   1st Qu.:175.8   1st Qu.:37.4    1st Qu.: 96.9
 Median :21.10    Median :47.00   Median : 81.70   Median :180.8   Median :40.0    Median :102.1
 Mean   :21.02    Mean   :46.98   Mean   : 82.69   Mean   :180.8   Mean   :40.1    Mean   :102.5
 3rd Qu.:27.20    3rd Qu.:55.00   3rd Qu.: 90.80   3rd Qu.:186.0   3rd Qu.:42.8    3rd Qu.:107.7
 Max.   :45.50    Max.   :81.00   Max.   :115.40   Max.   :200.5   Max.   :49.1    Max.   :124.7
  Abdomen_circ       Hip_Circ         Thin_Circ        Knee_Circ       Ankle_circ      Extended_Biceps_Circ
 Min.   : 70.40   Min.   : 86.5    Min.   :47.30    Min.   :31.70   Min.   :17.1     Min.   :22.80
 1st Qu.: 86.90   1st Qu.: 97.1    1st Qu.:57.50    1st Qu.:38.00   1st Qu.:22.7     1st Qu.:31.40
 Median : 93.60   Median :101.4    Median :61.30    Median :40.60   Median :25.0     Median :34.40
 Mean   : 94.07   Mean   :101.6    Mean   :61.39    Mean   :40.73   Mean   :25.2     Mean   :34.38
 3rd Qu.:100.90   3rd Qu.:106.0    3rd Qu.:65.00    3rd Qu.:43.40   3rd Qu.:27.7     3rd Qu.:37.20
 Max.   :120.10   Max.   :119.4    Max.   :76.20    Max.   :49.80   Max.   :32.9     Max.   :45.50
  Forearm_Circ      Wrist_Circ
 Min.   :19.00    Min.   :14.30
 1st Qu.:28.20    1st Qu.:18.00
 Median :30.70    Median :20.10
 Mean   :30.85    Mean   :20.43
 3rd Qu.:33.50    3rd Qu.:23.10
 Max.   :39.80    Max.   :27.90
```

```
> library(pastecs)
> stat.desc(Cholestrol)
```

	Body_FAT	Age	Weight	Height	Neck_Circ	Chest_Circ	Abdomen_circ
nbr.val	1429.0000000	1.429000e+03	1.429000e+03	1.429000e+03	1.429000e+03	1.429000e+03	1.429000e+03
nbr.null	0.0000000	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
nbr.na	0.0000000	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
min	-1.3000000	2.000000e+01	5.480000e+01	1.606000e+02	2.910000e+01	8.340000e+01	7.040000e+01
max	45.5000000	8.100000e+01	1.154000e+02	2.005000e+02	4.910000e+01	1.247000e+02	1.201000e+02
range	46.8000000	6.100000e+01	6.060000e+01	3.990000e+01	2.000000e+01	4.130000e+01	4.970000e+01
sum	30033.4000000	6.714100e+04	1.181593e+05	2.583725e+05	5.730570e+04	1.465289e+05	1.344327e+05
median	21.1000000	4.700000e+01	8.170000e+01	1.808000e+02	4.000000e+01	1.021000e+02	9.360000e+01
mean	21.0170749	4.698460e+01	8.268670e+01	1.808065e+02	4.010196e+01	1.025395e+02	9.407467e+01
SE.mean	0.2207919	3.385071e-01	3.089323e-01	1.871888e-01	9.628502e-02	2.096124e-01	2.571615e-01
CI.mean.0.95	0.4331113	6.640245e-01	6.060098e-01	3.671945e-01	1.888753e-01	4.111812e-01	5.044549e-01
var	69.6624113	1.637449e+02	1.363825e+02	5.007166e+01	1.324798e+01	6.278648e+01	9.450271e+01
std.dev	8.3464011	1.279628e+01	1.167829e+01	7.076133e+00	3.639777e+00	7.923792e+00	9.721250e+00
coef.var	0.3971248	2.723506e-01	1.412354e-01	3.913650e-02	9.076308e-02	7.727553e-02	1.033355e-01

	Hip_Circ	Thin_Circ	Knee_Circ	Ankle_circ	Extended_Biceps_Circ	Forearm_Circ
nbr.val	1.429000e+03	1.429000e+03	1.429000e+03	1.429000e+03	1.429000e+03	1.429000e+03
nbr.null	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
nbr.na	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
min	8.650000e+01	4.730000e+01	3.170000e+01	1.710000e+01	2.280000e+01	1.900000e+01
max	1.194000e+02	7.620000e+01	4.980000e+01	3.290000e+01	4.550000e+01	3.980000e+01
range	3.290000e+01	2.890000e+01	1.810000e+01	1.580000e+01	2.270000e+01	2.080000e+01
sum	1.452228e+05	8.772960e+04	5.819960e+04	3.601730e+04	4.912860e+04	4.408680e+04
median	1.014000e+02	6.130000e+01	4.060000e+01	2.500000e+01	3.440000e+01	3.070000e+01
mean	1.016255e+02	6.139230e+01	4.072750e+01	2.520455e+01	3.437971e+01	3.085150e+01
SE.mean	1.709364e-01	1.431641e-01	9.567779e-02	8.501066e-02	1.068870e-01	9.238317e-02
CI.mean.0.95	3.353134e-01	2.808344e-01	1.876841e-01	1.667592e-01	2.096723e-01	1.812213e-01
var	4.175431e+01	2.928871e+01	1.308141e+01	1.032712e+01	1.632607e+01	1.219601e+01
std.dev	6.461757e+00	5.411904e+00	3.616823e+00	3.213583e+00	4.040553e+00	3.492279e+00
coef.var	6.358403e-02	8.815282e-02	8.880542e-02	1.275001e-01	1.175273e-01	1.131964e-01

```
              Wrist_Circ
nbr.val       1.429000e+03
nbr.null      0.000000e+00
nbr.na        0.000000e+00
min           1.430000e+01
max           2.790000e+01
range         1.360000e+01
sum           2.919110e+04
median        2.010000e+01
mean          2.042764e+01
SE.mean       8.173729e-02
CI.mean.0.95  1.603380e-01
var           9.547128e+00
std.dev       3.089843e+00
coef.var      1.512579e-01
>
```

```
> library(moments)
> skewness(Cholestrol, na.rm = FALSE)
        Body_FAT                 Age              Weight              Height           Neck_Circ
     0.001257965         0.210584825         0.234682445         0.043445145         0.045222559
      Chest_Circ        Abdomen_circ            Hip_Circ           Thin_Circ           Knee_Circ
     0.263869618         0.173075851         0.158688485         0.114179108         0.055580073
      Ankle_circ  Extended_Biceps_Circ       Forearm_Circ          Wrist_Circ
     0.127468698         0.049552517         0.021894483         0.159851039
> kurtosis(Cholestrol, na.rm = FALSE)
        Body_FAT                 Age              Weight              Height           Neck_Circ
        2.403113            2.528057            2.564084            2.569537            2.424320
      Chest_Circ        Abdomen_circ            Hip_Circ           Thin_Circ           Knee_Circ
        2.683336            2.487484            2.665441            2.679735            2.319728
      Ankle_circ  Extended_Biceps_Circ       Forearm_Circ          Wrist_Circ
        2.109038            2.541955            2.399816            1.984189
>
```

```
> cov(Cholestrol, y = NULL, use = "all.obs",
+     method = c("pearson", "kendall", "spearman"))
```
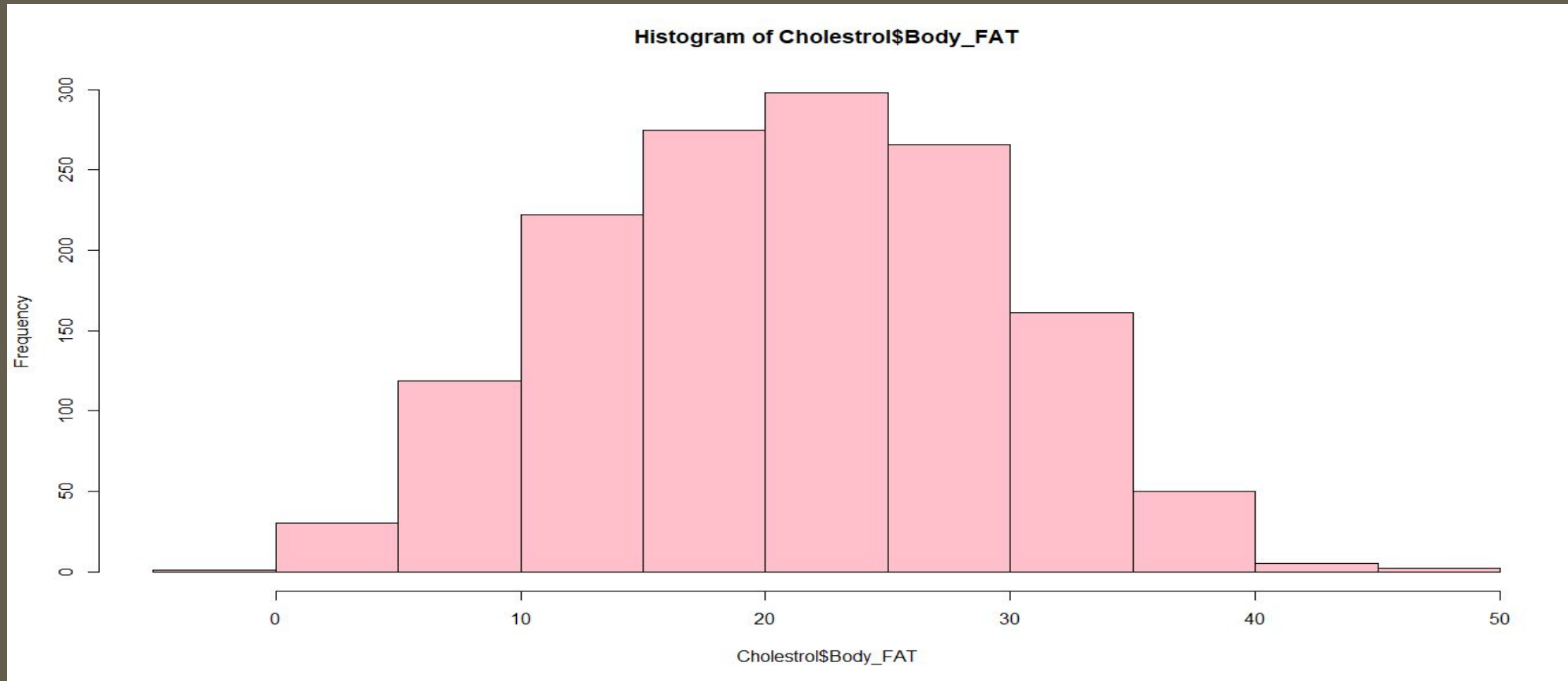
	Body_FAT	Age	Weight	Height	Neck_Circ	Chest_Circ	Abdomen_circ	Hip_Circ
Body_FAT	69.662411	36.108106	59.333434	6.911906	15.437460	46.34525	66.53246	35.066007
Age	36.108106	163.744861	3.118072	-11.184774	12.054722	24.57932	36.06698	1.757115
Weight	59.333434	3.118072	136.382540	46.757377	26.245691	81.23160	98.29333	68.400626
Height	6.911906	-11.184774	46.757377	50.071659	12.841983	19.10038	19.79617	23.844120
Neck_Circ	15.437460	12.054722	26.245691	12.841983	13.247979	19.60746	21.48174	16.146638
Chest_Circ	46.345250	24.579320	81.231597	19.100380	19.607465	62.78648	69.33849	41.722040
Abdomen_circ	66.532464	36.066977	98.293332	19.796173	21.481744	69.33849	94.50271	52.854063
Hip_Circ	35.066007	1.757115	68.400626	23.844120	16.146638	41.72204	52.85406	41.754309
Thin_Circ	26.825748	-4.968116	50.766228	18.335127	14.403264	30.65257	38.06255	31.371331
Knee_Circ	16.083865	8.747272	27.772061	15.283581	11.322859	18.71988	21.79654	18.211785
Ankle_circ	10.208816	6.000910	16.512266	12.317043	9.854004	12.18828	12.59440	12.156621
Extended_Biceps_Circ	17.737195	6.579939	31.418420	13.927520	12.592365	21.95085	23.81752	19.220531
Forearm_Circ	12.828602	6.628244	21.308956	12.263838	11.179325	15.80143	16.23342	14.151509
Wrist_Circ	9.731580	11.241602	13.250515	10.601473	9.933951	11.30178	11.55036	10.445241

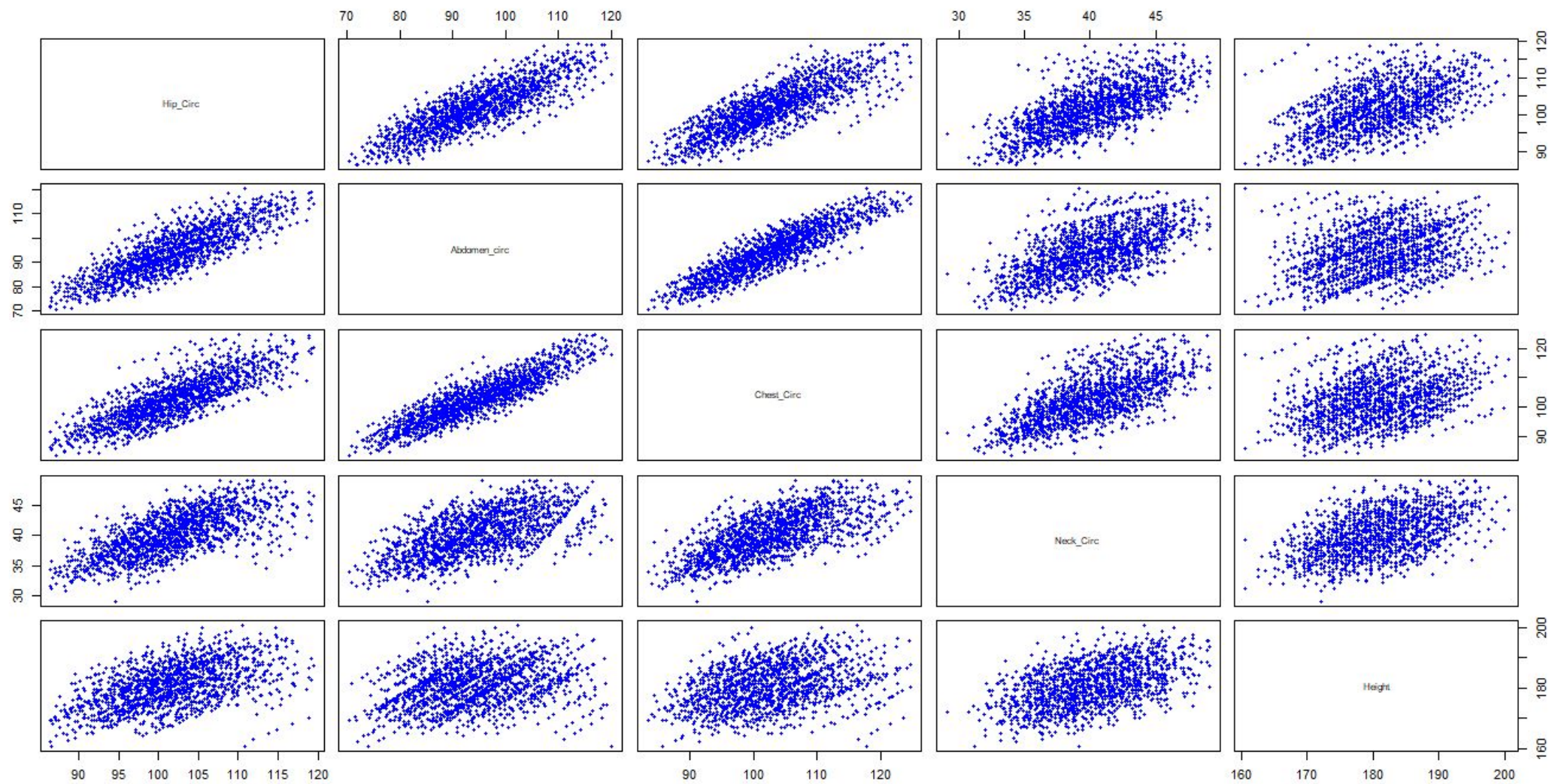|	Thin_Circ	Knee_Circ	Ankle_circ	Extended_Biceps_Circ	Forearm_Circ	Wrist_Circ
Body_FAT	26.825748	16.083865	10.208816	17.737195	12.828602	9.731580
Age	-4.968116	8.747272	6.000910	6.579939	6.628244	11.241602
Weight	50.766228	27.772061	16.512266	31.418420	21.308956	13.250515
Height	18.335127	15.283581	12.317043	13.927520	12.263838	10.601473
Neck_Circ	14.403264	11.322859	9.854004	12.592365	11.179325	9.933951
Chest_Circ	30.652566	18.719880	12.188280	21.950851	15.801432	11.301779
Abdomen_circ	38.062550	21.796545	12.594401	23.817518	16.233420	11.550358
Hip_Circ	31.371331	18.211785	12.156621	19.220531	14.151509	10.445241
Thin_Circ	29.288708	15.846872	11.660651	17.456440	13.209647	10.048224
Knee_Circ	15.846872	13.081407	10.374910	12.153864	10.734416	9.683238
Ankle_circ	11.660651	10.374910	10.327115	10.218139	9.787784	9.360680
Extended_Biceps_Circ	17.456440	12.153864	10.218139	16.326072	12.178616	9.927949
Forearm_Circ	13.209647	10.734416	9.787784	12.178616	12.196015	9.592826
Wrist_Circ	10.048224	9.683238	9.360680	9.927949	9.592826	9.547128

```
> cor(Cholestrol, y = NULL, use = "all.obs",
+     method = c("pearson", "kendall", "spearman"))
                      Body_FAT         Age     Weight     Height Neck_Circ Chest_Circ Abdomen_circ
Body_FAT             1.0000000  0.33808167 0.60872459  0.1170314 0.5081615  0.7007658    0.8199969
Age                  0.3380817  1.00000000 0.02086522 -0.1235229 0.2588204  0.2424113    0.2899371
Weight               0.6087246  0.02086522 1.00000000  0.5658154 0.6174529  0.8778343    0.8658099
Height               0.1170314 -0.12352287 0.56581539  1.0000000 0.4986103  0.3406536    0.2877817
Neck_Circ            0.5081615  0.25882042 0.61745286  0.4986103 1.0000000  0.6798507    0.6071173
Chest_Circ           0.7007658  0.24241135 0.87783427  0.3406536 0.6798507  1.0000000    0.9001590
Abdomen_circ         0.8199969  0.28993709 0.86580991  0.2877817 0.6071173  0.9001590    1.0000000
Hip_Circ             0.6501842  0.02125033 0.90642117  0.5214764 0.6865254  0.8148578    0.8414060
Thin_Circ            0.5938852 -0.07173940 0.80324014  0.4787820 0.7311997  0.7147985    0.7234785
Knee_Circ            0.5327997  0.18899990 0.65750868  0.5971754 0.8601101  0.6531949    0.6199238
Ankle_circ           0.3806156  0.14592974 0.43998492  0.5416528 0.8424583  0.4786520    0.4031493
Extended_Biceps_Circ 0.5259505  0.12726155 0.66583114  0.4871211 0.8562323  0.6856104    0.6063642
Forearm_Circ         0.4401200  0.14832205 0.52248497  0.4962739 0.8794918  0.5710241    0.4781663
Wrist_Circ           0.3773529  0.28432040 0.36721212  0.4848795 0.8833053  0.4616123    0.3845359
                       Hip_Circ  Thin_Circ Knee_Circ Ankle_circ Extended_Biceps_Circ Forearm_Circ Wrist_Circ
Body_FAT             0.65018421  0.5938852 0.5327997  0.3806156            0.5259505    0.4401200  0.3773529
Age                  0.02125033 -0.0717394 0.1889999  0.1459297            0.1272615    0.1483220  0.2843204
Weight               0.90642117  0.8032401 0.6575087  0.4399849            0.6658311    0.5224850  0.3672121
Height               0.52147640  0.4787820 0.5971754  0.5416528            0.4871211    0.4962739  0.4848795
Neck_Circ            0.68652543  0.7311997 0.8601101  0.8424583            0.8562323    0.8794918  0.8833053
Chest_Circ           0.81485778  0.7147985 0.6531949  0.4786520            0.6856104    0.5710241  0.4616123
Abdomen_circ         0.84140601  0.7234785 0.6199238  0.4031493            0.6063642    0.4781663  0.3845359
Hip_Circ             1.00000000  0.8970821 0.7792461  0.5854269            0.7361628    0.6271092  0.5231563
Thin_Circ            0.89708208  1.0000000 0.8095920  0.6704759            0.7982974    0.6989275  0.6009009
Knee_Circ            0.77924610  0.8095920 1.0000000  0.8926220            0.8316609    0.8498499  0.8664769
Ankle_circ           0.58542695  0.6704759 0.8926220  1.0000000            0.7869396    0.8721394  0.9427174
Extended_Biceps_Circ 0.73616282  0.7982974 0.8316609  0.7869396            1.0000000    0.8630741  0.7952109
Forearm_Circ         0.62710915  0.6989275 0.8498499  0.8721394            0.8630741    1.0000000  0.8889989
Wrist_Circ           0.52315628  0.6009009 0.8664769  0.9427174            0.7952109    0.8889989  1.0000000
> 
```
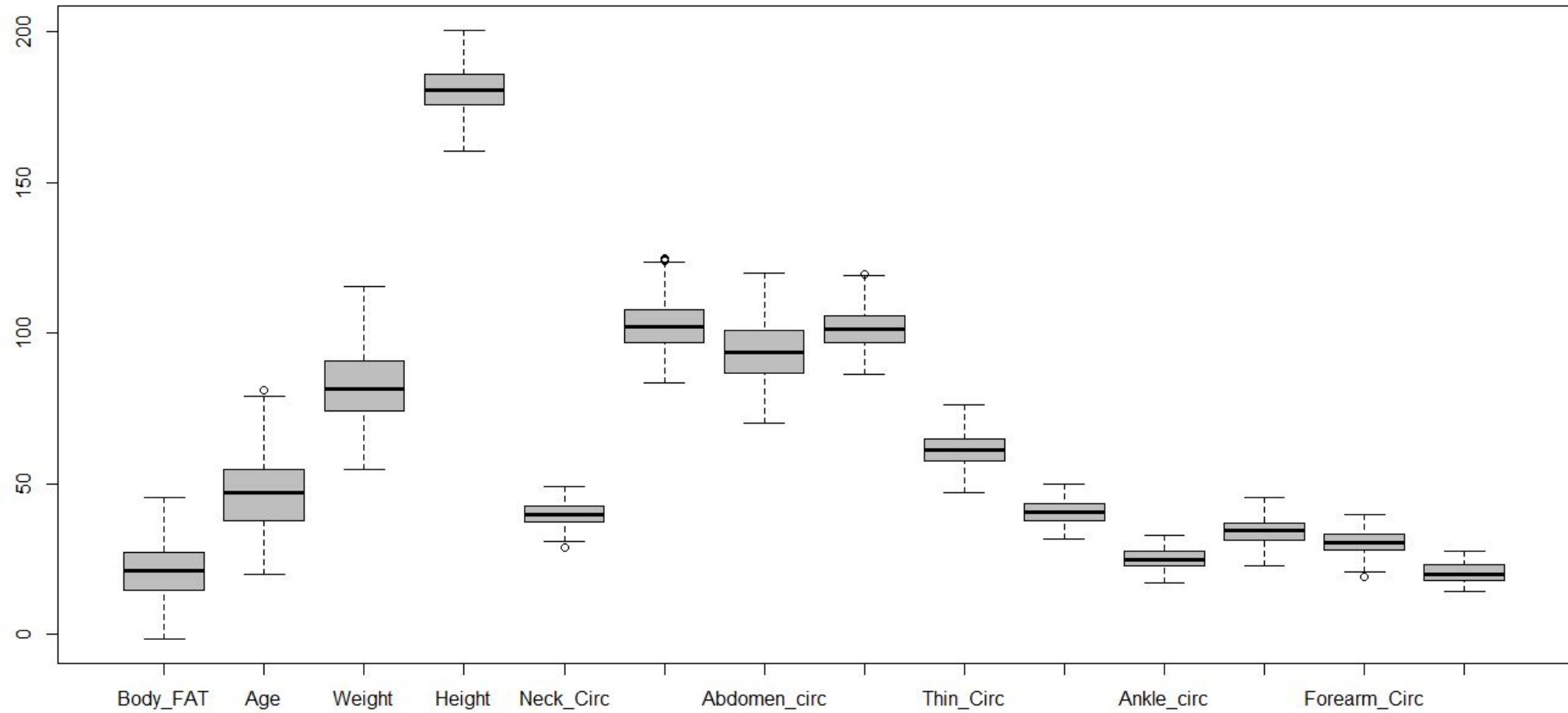
VISUALIZATION OF DATA

Histogram of Cholestrol$Body_FAT

boxplot(*, horizontal = FALSE)

```
> chol = lm(Body_FAT~. , data = Cholestrol)
> summary(chol)

Call:
lm(formula = Body_FAT ~ ., data = Cholestrol)

Residuals:
     Min       1Q   Median       3Q      Max
-10.7086  -3.1196   0.0612   3.2197  10.3771

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)          -55.27544    7.20976  -7.667 3.26e-14 ***
Age                    0.04396    0.01322   3.326 0.000905 ***
Weight                -0.37569    0.04658  -8.065 1.55e-15 ***
Height                -0.01179    0.03052  -0.386 0.699356
Neck_Circ             -0.31098    0.09801  -3.173 0.001542 **
Chest_Circ            -0.03006    0.04438  -0.677 0.498308
Abdomen_circ           0.97669    0.03685  26.504  < 2e-16 ***
Hip_Circ              -0.03162    0.06274  -0.504 0.614326
Thin_Circ              0.24595    0.06290   3.910 9.66e-05 ***
Knee_Circ              0.07676    0.11100   0.692 0.489328
Ankle_circ             0.31408    0.13047   2.407 0.016200 *
Extended_Biceps_Circ   0.24120    0.07192   3.354 0.000819 ***
Forearm_Circ           0.34636    0.09028   3.836 0.000130 ***
Wrist_Circ            -0.52929    0.18787  -2.817 0.004910 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.281 on 1415 degrees of freedom
Multiple R-squared:  0.7393,    Adjusted R-squared:  0.7369
F-statistic: 308.7 on 13 and 1415 DF,  p-value: < 2.2e-16
```

```
> chol1 = lm(Body_FAT~Age + Weight + Abdomen_circ+ Thin_Circ + Extended_Biceps_Circ + Forearm_Circ, data = Choles
trol)
> summary(chol1)

Call:
lm(formula = Body_FAT ~ Age + Weight + Abdomen_circ + Thin_Circ +
    Extended_Biceps_Circ + Forearm_Circ, data = Cholestrol)

Residuals:
     Min      1Q  Median      3Q     Max
-10.8884  -3.0709  -0.0217   3.2473  10.2024

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)          -58.80169    1.57698 -37.288  < 2e-16 ***
Age                    0.01679    0.01186   1.416   0.1569
Weight                -0.37945    0.02461 -15.418  < 2e-16 ***
Abdomen_circ           0.95796    0.02936  32.633  < 2e-16 ***
Thin_Circ              0.23538    0.04910   4.794 1.81e-06 ***
Extended_Biceps_Circ   0.11613    0.06874   1.689   0.0914 .
Forearm_Circ           0.05973    0.06719   0.889   0.3741
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.333 on 1422 degrees of freedom
Multiple R-squared:  0.7316,    Adjusted R-squared:  0.7305
F-statistic: 646.1 on 6 and 1422 DF,  p-value: < 2.2e-16
```

```
> chol2 = lm(Body_FAT~Weight + Abdomen_circ + Thin_Circ, data = Cholestrol)
> summary(chol2)

Call:
lm(formula = Body_FAT ~ Weight + Abdomen_circ + Thin_Circ, data = Cholestrol)

Residuals:
     Min       1Q   Median       3Q      Max
-10.6307  -3.0667   0.0495   3.2598  10.6613

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -58.60093    1.54691 -37.883   <2e-16 ***
Weight        -0.39254    0.02297 -17.092   <2e-16 ***
Abdomen_circ   0.98489    0.02381  41.371   <2e-16 ***
Thin_Circ      0.31637    0.03592   8.807   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.356 on 1425 degrees of freedom
Multiple R-squared:  0.7281,    Adjusted R-squared:  0.7276
F-statistic:  1272 on 3 and 1425 DF,  p-value: < 2.2e-16
```

```
> chol3= lm(Body_FAT~Weight + Abdomen_circ + Thin_Circ+ Neck_Circ + Wrist_Circ , data = Cholestrol)
> summary(chol3)

Call:
lm(formula = Body_FAT ~ Weight + Abdomen_circ + Thin_Circ + Neck_Circ +
    Wrist_Circ, data = Cholestrol)

Residuals:
    Min      1Q   Median      3Q      Max
-10.1665  -3.0988   0.0835   3.1766  11.0329

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -55.98583    1.90321 -29.416  < 2e-16 ***
Weight        -0.36872    0.02480 -14.870  < 2e-16 ***
Abdomen_circ   0.98782    0.02411  40.970  < 2e-16 ***
Thin_Circ      0.28548    0.04284   6.664  3.8e-11 ***
Neck_Circ     -0.19575    0.09188  -2.131   0.0333 *
Wrist_Circ     0.23920    0.09430   2.537   0.0113 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.35 on 1423 degrees of freedom
Multiple R-squared:  0.7294,    Adjusted R-squared:  0.7284
F-statistic:   767 on 5 and 1423 DF,  p-value: < 2.2e-16

>
```

THANK YOU