STAT 463
Assignment 1
Xin Gao (43044879)

Question 1

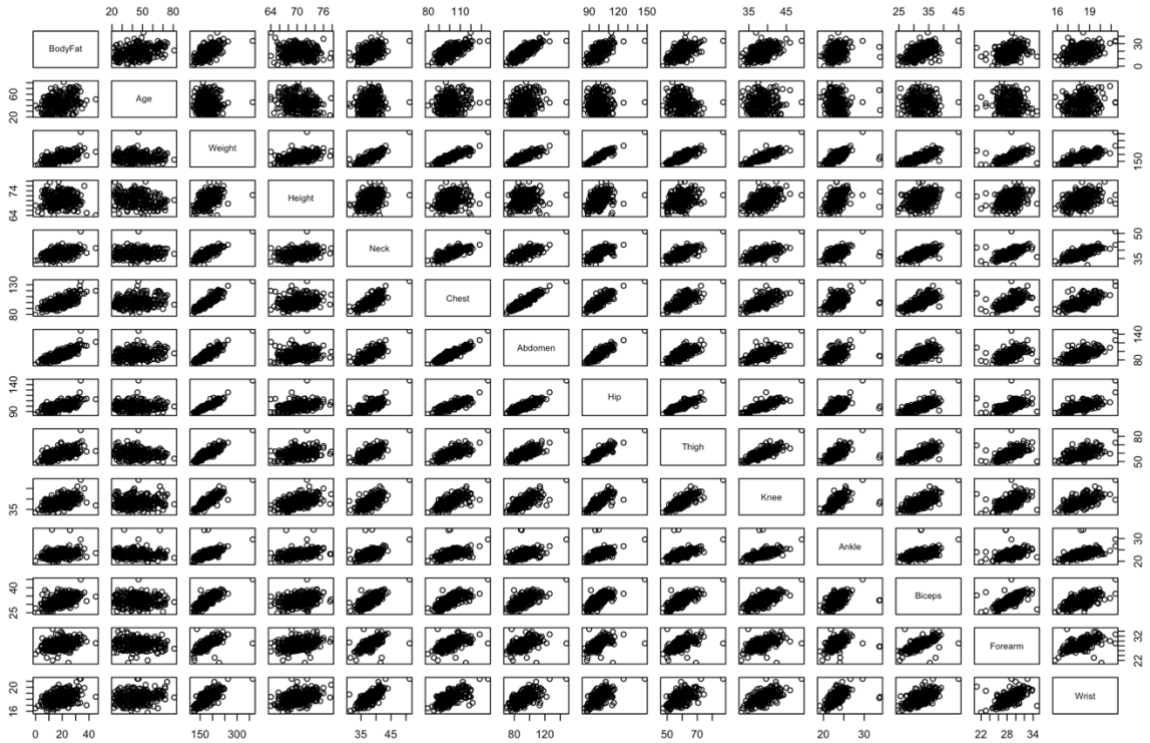
a. # Find out what measures are well correlated with body fat:

```
> round(cor(bodyfat[, -1]), 2)
```

	BodyFat	Age	Weight	Height	Neck	Chest	Abdomen	Hip	Thigh	Knee	Ankle	Biceps	Forearm	Wrist
BodyFat	1.00	0.29	0.61	-0.02	0.49	0.70	0.81	0.63	0.56	0.51	0.27	0.49	0.36	0.35
Age	0.29	1.00	-0.01	-0.25	0.11	0.18	0.23	-0.05	-0.20	0.02	-0.11	-0.04	-0.09	0.21
Weight	0.61	-0.01	1.00	0.49	0.83	0.89	0.89	0.94	0.87	0.85	0.61	0.80	0.63	0.73
Height	-0.02	-0.25	0.49	1.00	0.32	0.23	0.19	0.37	0.34	0.50	0.39	0.32	0.32	0.40
Neck	0.49	0.11	0.83	0.32	1.00	0.78	0.75	0.73	0.70	0.67	0.48	0.73	0.62	0.74
Chest	0.70	0.18	0.89	0.23	0.78	1.00	0.92	0.83	0.73	0.72	0.48	0.73	0.58	0.66
Abdomen	0.81	0.23	0.89	0.19	0.75	0.92	1.00	0.87	0.77	0.74	0.45	0.68	0.50	0.62
Hip	0.63	-0.05	0.94	0.37	0.73	0.83	0.87	1.00	0.90	0.82	0.56	0.74	0.55	0.63
Thigh	0.56	-0.20	0.87	0.34	0.70	0.73	0.77	0.90	1.00	0.80	0.54	0.76	0.57	0.56
Knee	0.51	0.02	0.85	0.50	0.67	0.72	0.74	0.82	0.80	1.00	0.61	0.68	0.56	0.66
Ankle	0.27	-0.11	0.61	0.39	0.48	0.48	0.45	0.56	0.54	0.61	1.00	0.48	0.42	0.57
Biceps	0.49	-0.04	0.80	0.32	0.73	0.73	0.68	0.74	0.76	0.68	0.48	1.00	0.68	0.63
Forearm	0.36	-0.09	0.63	0.32	0.62	0.58	0.50	0.55	0.57	0.56	0.42	0.68	1.00	0.59
Wrist	0.35	0.21	0.73	0.40	0.74	0.66	0.62	0.63	0.56	0.66	0.57	0.63	0.59	1.00

Matrix plot – Person factor left out

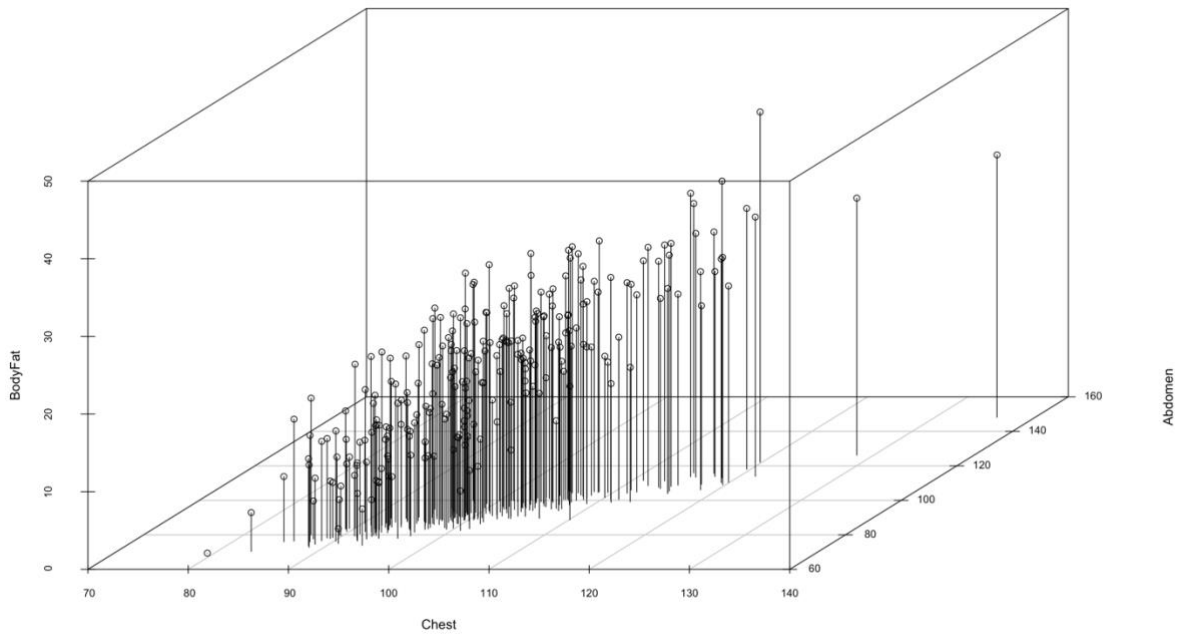
```
> pairs(bodyfat[, -1])
```



b. The two most useful measures of body fat : ***Chest & Abdomen***

c. #Creat a sensible 3-D plot with “Chest”, “Abdomen” and “BodyFat”

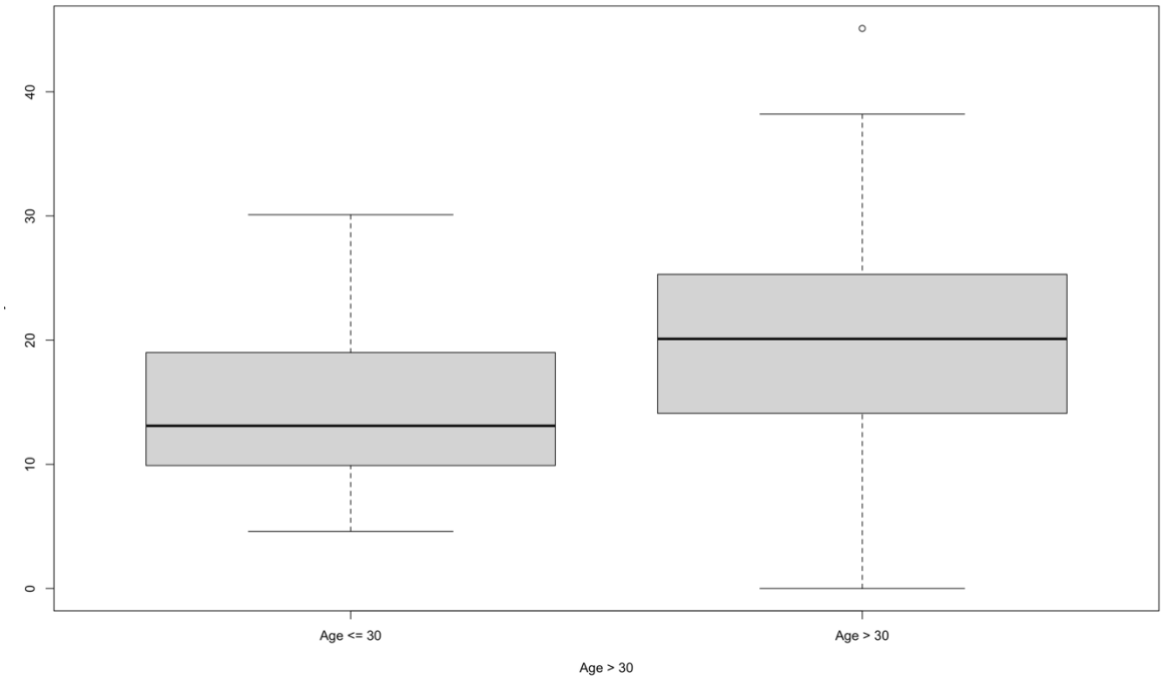
```
> library(scatterplot3d)
> with(bodyfat, scatterplot3d(Chest, Abdomen, BodyFat, type = "h"))
```



"Chest", "Abdomen" and "BodyFat" are factors that are quite well linear correlated.

d. # Group the data into two age-groups(<= 30 & >30), compare body fat between the two age-groups

```
> with(bodyfat, boxplot(BodyFat ~ Age > 30, names = c("Age <= 30", "Age > 30")))
```



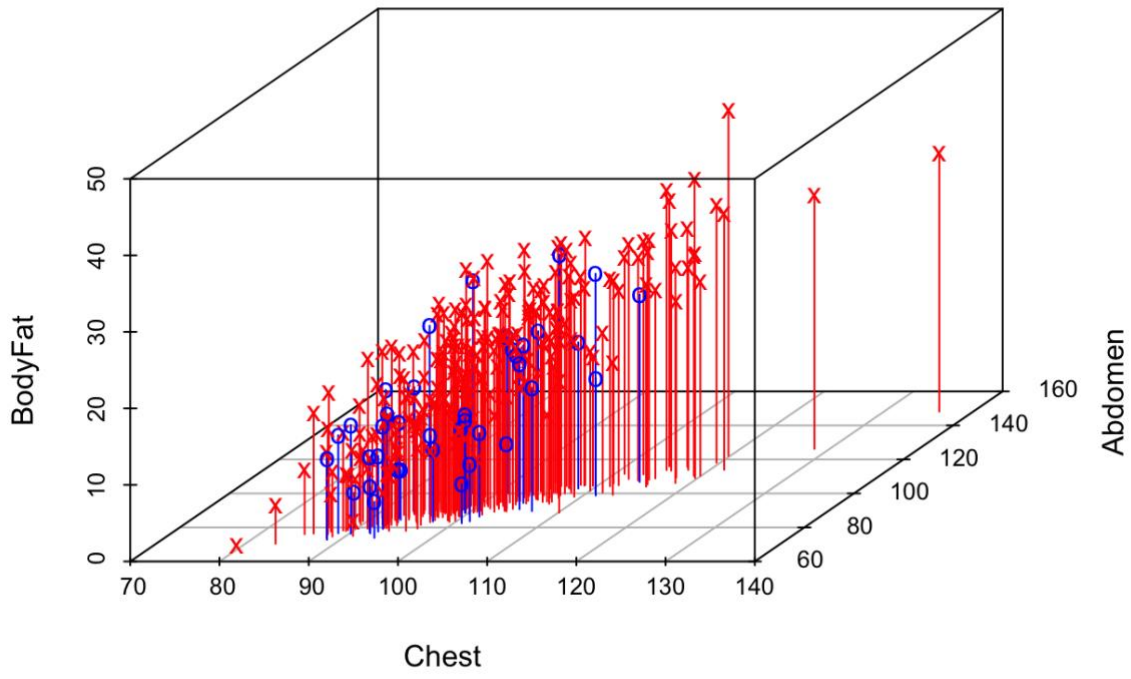
Compared to people aged more than 30, people aged less than 30 are with lower body fat.

e. #Repeat the analysis for the two groups separately

```
> round(cor(bodyfat[bodyfat$Age > 30, -1]), 2)
        BodyFat   Age Weight Height Neck Chest Abdomen   Hip Thigh Knee Ankle Biceps Forearm Wrist
BodyFat    1.00  0.20   0.65   0.04 0.53  0.72    0.81  0.67  0.61 0.51  0.29   0.53    0.43  0.37
Age        0.20  1.00  -0.02  -0.22 0.09  0.15    0.17 -0.04 -0.20 0.03 -0.06  -0.04   -0.04  0.27
Weight     0.65 -0.02   1.00   0.50 0.84  0.91    0.91  0.94  0.87 0.85  0.60   0.80    0.65  0.72
Height     0.04 -0.22   0.50   1.00 0.35  0.27    0.24  0.38  0.34 0.54  0.38   0.33    0.34  0.40
Neck       0.53  0.09   0.84   0.35 1.00  0.79    0.77  0.74  0.71 0.68  0.48   0.74    0.63  0.74
Chest      0.72  0.15   0.91   0.27 0.79  1.00    0.93  0.84  0.74 0.72  0.49   0.74    0.62  0.66
Abdomen    0.81  0.17   0.91   0.24 0.77  0.93    1.00  0.89  0.79 0.73  0.46   0.70    0.55  0.63
Hip        0.67 -0.04   0.94   0.38 0.74  0.84    0.89  1.00  0.89 0.82  0.54   0.73    0.56  0.62
Thigh      0.61 -0.20   0.87   0.34 0.71  0.74    0.79  0.89  1.00 0.79  0.51   0.75    0.58  0.54
Knee       0.51  0.03   0.85   0.54 0.68  0.72    0.73  0.82  0.79 1.00  0.60   0.67    0.59  0.67
Ankle      0.29 -0.06   0.60   0.38 0.48  0.49    0.46  0.54  0.51 0.60  1.00   0.48    0.43  0.56
Biceps     0.53 -0.04   0.80   0.33 0.74  0.74    0.70  0.73  0.75 0.67  0.48   1.00    0.71  0.63
Forearm    0.43 -0.04   0.65   0.34 0.63  0.62    0.55  0.56  0.58 0.59  0.43   0.71    1.00  0.59
Wrist      0.37  0.27   0.72   0.40 0.74  0.66    0.63  0.62  0.54 0.67  0.56   0.63    0.59  1.00
```

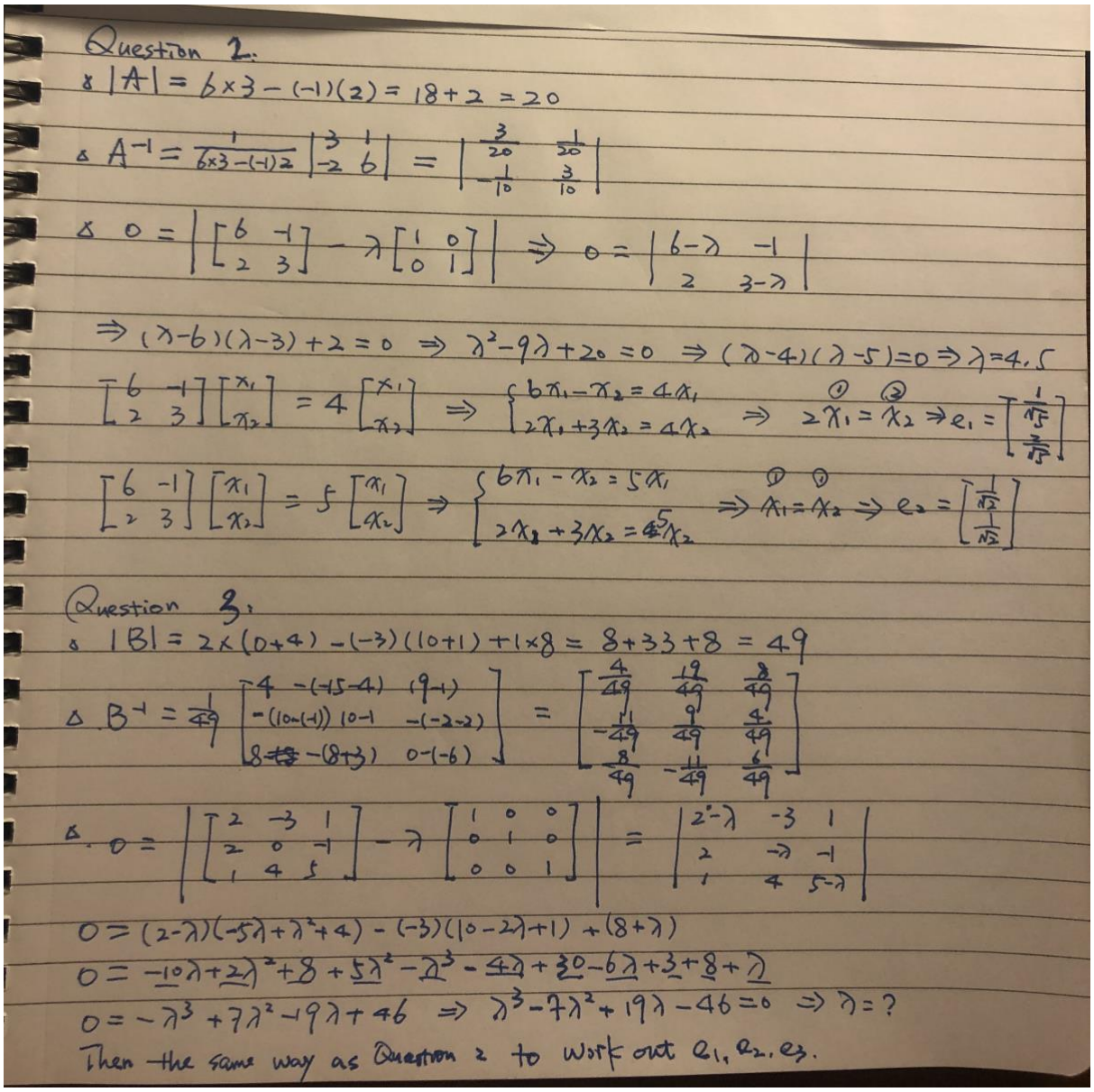
The two most useful measures of body fat are still : ***Chest & Abdomen***

```
> with(bodyfat, scatterplot3d(Chest, Abdomen, BodyFat, type = "h", color = ifelse(Age > 30, "red", "blu
e"), pch = ifelse(Age <= 30, "o", "x")))
```



"Chest", "Abdomen" and "BodyFat" are factors that are quite well linear correlated. People aged more than 30(red) are with higher body fat while people aged below 30 have less body fat but more variability.

Question 2 & 3


Question 2:
Question 3:
Then the same way as Question 2 to work out e1, e2, e3.