Exercise 1

The task:

a) Calculate the estimated variance of the sum score $X_1 + X_2 + X_3 + X_4$.

10,66

b) Calculate the estimated correlation matrix from the above covariance matrix.

Covariance

	x1	x2	x3	x4
x1	0,25	0,12	0,36	0,16
x2	0,12	0,77	0,66	0,35
x3	0,36	0,66	3,68	0,74
x4	0,16	0,35	0,74	1,18

Std. Deviation (Square root of Covariance)

	x1	x2	x3	x4
x1	0,50	0,35	0,60	0,40
x2	0,35	0,88	0,81	0,59
x3	0,60	0,81	1,92	0,86
x4	0,40	0,59	0,86	1,09

Corelation (The covariance of two variables divided by the product of their standard deviation)

	x1	x2	x3	x4
x1	1,00	0,69	1,20	0,80
x2	0,69	1,00	0,93	0,67
x3	1,20	0,93	1,00	0,45
x4	0,80	0,67	0,45	1,00

c) Comment on the relationships between the variables X1, X2, X3 and X4 based on the observations.

The matrix is symmetric since the upper and lower diagonals contain identical entries.

d) What further information would you like to have in order better understand the relationships between the variables X1, X2, X3 and X4?

The number of observations for each of the variable would have provided us with more contextual information. In addition, it would be interesting to know the spread of these observations to observe if there are many widely different values.

Exercise 2

(0,3*0,7)+ (0,2*0,8)+ (0,3*0,7)+ (0,94*0,06)+ (0,69*0,31)+ (0,52*0,48)+ (0,69*0,31)+ (0,52*0,48)/3,67
= 1,353/3,67
=0,368

Exercise 3

a)

The response options can be scored for never being 0, and increasing unit of one for every successive response option as they are ordered.

Every day = 4
At least once a week but not every day = 3
Less than once a week but at least once a month = 2
Less than once a month = 1
Never = 0

The response options can be scored for never being 0, and increasing unit of one for every successive response option as they are ordered on a spectrum of 0 never heard to 3 indicating familiarity

I am familiar with this and I would be able to explain this well = 3
I know something about this and could explain the general issue = 2
I have heard about this but I would not be able to explain what it is really about = 1
I have never heard of this = 0

These are nominal variables and the categories cannot be ordered, so even the scoring is also nominal

Denmark = 5
Finland = 4
Norway = 3

Sweden = 2
Iceland = 1

As there are only binary values in Gender, so 1 and 0 were used as scoring

Male = 1
Female = 0

These are nominal variables and the categories cannot be ordered, so even the scoring is also nominal

Yellow = 4
Red = 3
Blue = 2
Green = 1

b)

```
> summary(QT1$X1..Use.email.)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  3.000   5.000   5.000   4.765   5.000   5.000 
> summary(QT1$X3..Nuclear.waste.)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  1.000   2.000   2.000   1.765   2.000   2.000 
> summary(QT1$what.is.your.gender.)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
 0.0000  0.0000  1.0000  0.6471  1.0000  1.0000 
```

c)

The objective of the questionnaire seems to be to score respondents on their usage of media and their awareness of environment. Therefore, scoring ranges to indicate low, medium, high media usage or environmental awareness seems to make sense. The scoring is obtained by taking the highest score possible in each of the parts as the upper limit and vice-versa and dividing into three equal parts.

Part 1

High -24-35

Moderate – 12-23

Low - 0-12

Part 2

High -15-21

Moderate – 8-14

Low Awareness- 0-7

6)

Respondent	Part 1	Part 2
1	22	14
2	23	13
3	23	12
4	23	11
5	17	13
6	21	10
7	25	16
8	20	15
9	17	19
10	21	7
11	21	14
12	22	15
13	16	17
14	16	13
15	26	15
16	23	15
17	19	15