Chapter 1 practice problems

1.1

	x_1	x_2	x_1^2	x_2^2	x_1x_2
	3	5	9	25	15
	4	5.5	16	30.25	22
	2	4	4	16	8
	6	7	36	49	42
	8	10	64	100	80
	2	5	4	25	10
	5	7.5	25	56.25	37.5
Sum	30	44	158	301.5	214.5
Mean	4.2857	6.2857			

$$\sum_j (x_{j1} - \bar{x}_1)^2 = \sum_j x_{j1}^2 - \frac{\left(\sum_j x_{j1}\right)^2}{n} = 158 - \frac{30^2}{7} = 29.4286$$

$$\sum_j (x_{j2} - \bar{x}_2)^2 = \sum_j x_{j2}^2 - \frac{\left(\sum_j x_{j2}\right)^2}{n} = 301.5 - \frac{44^2}{7} = 24.9286$$

$$\sum_j (x_{j1} - \bar{x}_1)(x_{j2} - \bar{x}_2) = \sum_j x_{j1}x_{j2} - \frac{\left(\sum_j x_{j1}\right)\left(\sum_j x_{j2}\right)}{n} = 214.5 - \frac{(30)(44)}{7} = 25.9286$$

$$s_{11} = \frac{29.4286}{7} = 4.2041 \quad s_{22} = \frac{24.9286}{7} = 3.5612 \quad s_{12} = \frac{25.9286}{7} = 3.7041$$

1.2

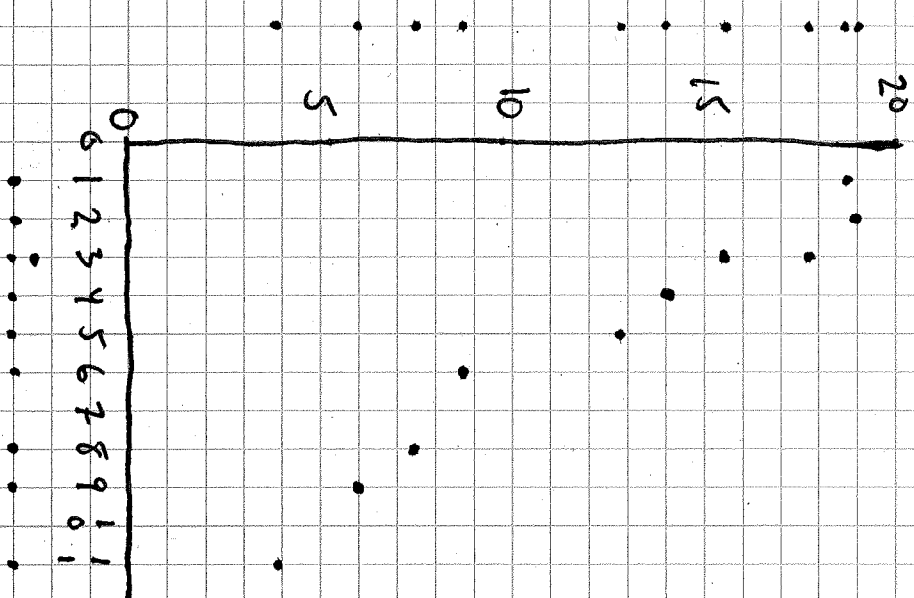
X_1	X_2	X_1^2	X_2^2	X_1X_2
1	18.95	1	359.1025	18.95
2	19.00	4	361.0000	38.00
3	17.95	9	322.2025	53.85
3	15.54	9	241.4916	46.62
4	14.00	16	196.0000	56.00
5	12.95	25	167.7025	64.75
6	8.94	36	79.9326	53.64
8	7.49	64	56.1001	59.92
9	6.00	81	36.0000	54.00
11	3.99	121	15.9201	43.89
Sum 52	124.81	366	1835.4519	489.62
Mean 5.2	12.481			

$$S_{11} = \frac{1}{10}\left[366 - \frac{52^2}{10}\right] = \frac{1}{10}(95.6) = 9.56$$

$$S_{22} = \frac{1}{10}\left[1835.4519 - \frac{(124.81)^2}{10}\right] = \frac{1}{10}(277.6983) = 27.7698$$

$$S_{12} = \frac{1}{10}\left[489.62 - \frac{(52)(124.81)}{10}\right] = \frac{1}{10}(-159.392) = -15.9392$$

$$r_{12} = \frac{-15.9392}{\sqrt{9.56}\,\sqrt{27.7698}} = \frac{-15.9392}{16.2935} = -.9783$$



$$\bar{X} = \begin{bmatrix} 5.2 \\ 12.481 \end{bmatrix}$$

$$S_n = \begin{bmatrix} 9.56 & -15.9392 \\ -15.9392 & 27.7698 \end{bmatrix}$$

$$R = \begin{bmatrix} 1 & -.9783 \\ -.9783 & 1 \end{bmatrix}$$

1.3

X_1	X_2	X_3	X_1^2	X_2^2	X_3^2	X_1X_2	X_1X_3	X_2X_3
9	12	3	81	144	9	108	27	36
2	8	4	4	64	16	16	8	32
6	6	0	36	36	0	36	0	0
5	4	2	25	16	4	20	10	8
8	10	1	64	100	1	80	8	10
30	40	10	210	360	30	270	63	86
Sum								
Mean	6	8	2					

$$s_{11} = \frac{1}{5}\left[210 - \frac{30^2}{5}\right] = \frac{1}{5}(30) = 6 \qquad s_{12} = \frac{1}{5}\left[270 - \frac{30(40)}{5}\right] = 6$$

$$s_{22} = \frac{1}{5}\left[360 - \frac{40^2}{5}\right] = \frac{1}{5}(40) = 8 \qquad s_{13} = \frac{1}{5}\left[63 - \frac{30(10)}{5}\right] = 0.6$$

$$s_{33} = \frac{1}{5}\left(30 - \frac{10^2}{5}\right) = \frac{1}{5}(10) = 2 \qquad s_{23} = \frac{1}{5}\left[86 - \frac{40(10)}{5}\right] = 1.2$$

$$r_{12} = \frac{6}{\sqrt{6}\sqrt{8}} = \frac{6}{6.9282} = 0.8660$$

$$r_{13} = \frac{0.6}{\sqrt{6}\sqrt{2}} = \frac{0.6}{3.4641} = 0.1732$$

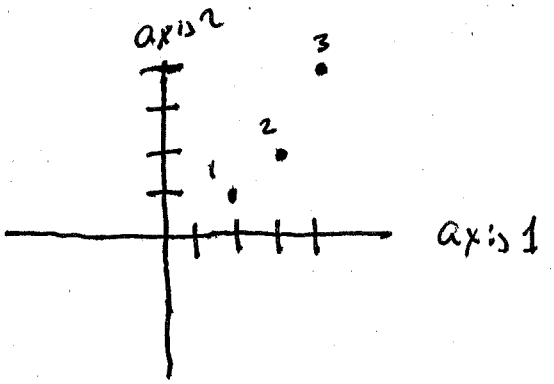
$$r_{23} = \frac{1.2}{\sqrt{8}\sqrt{2}} = \frac{1.2}{4} = 0.3000$$

$$\bar{X} = \begin{bmatrix} 6 \\ 8 \\ 2 \end{bmatrix}$$

$$S_n = \begin{bmatrix} 6 & 6 & 0.6 \\ 6 & 8 & 1.2 \\ 0.6 & 1.2 & 2 \end{bmatrix}$$

$$R = \begin{bmatrix} 1 & .8660 & .1732 \\ .8660 & 1 & .3000 \\ .1732 & .3000 & 1 \end{bmatrix}$$

1.7 a) (2, 1), (3, 2), (4, 4)



Problem 1.6

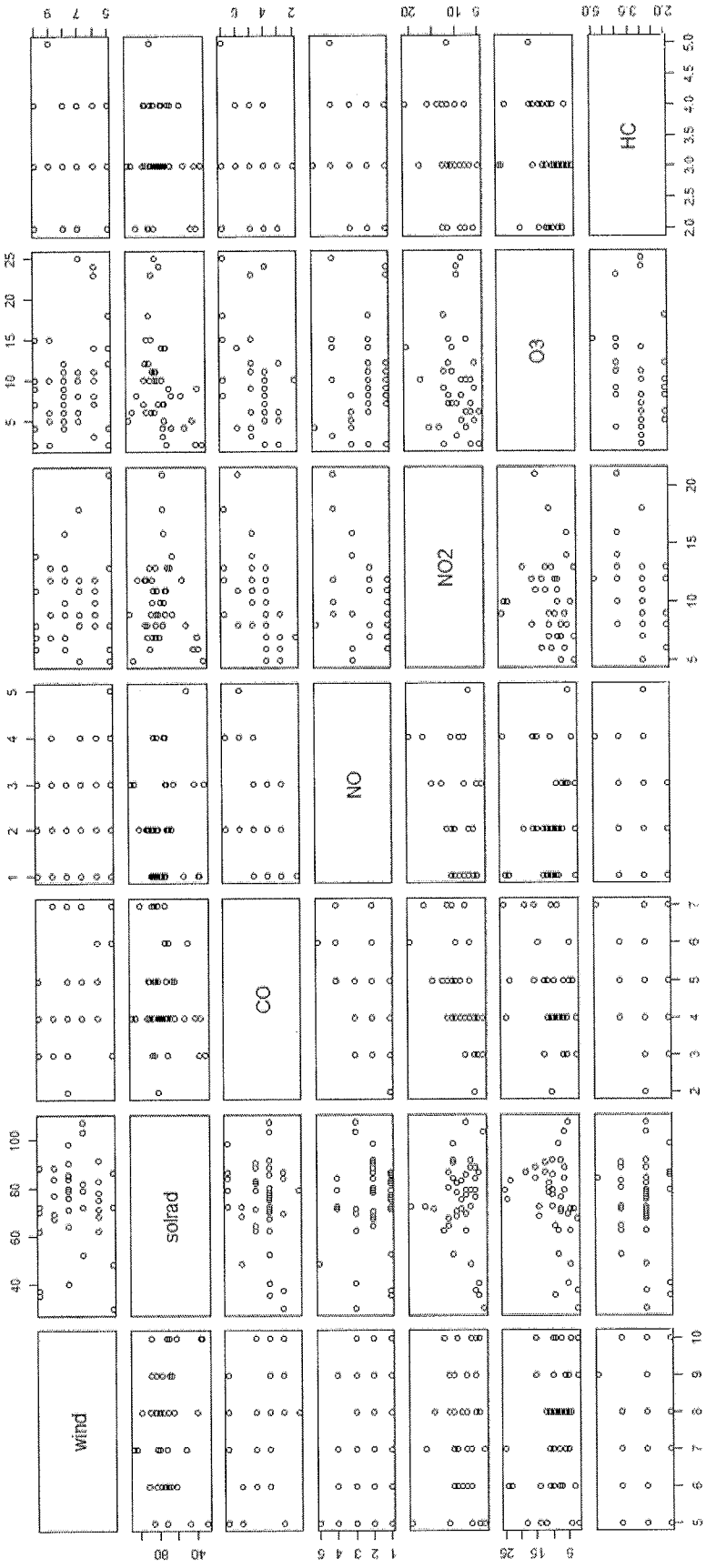
R Program (Constructs scatterplot matrix – not dot plots)

```
prob1.6 <-
read.table("http://www.stat.ufl.edu/~winner/sta4702/data/wichern/T1-5.dat",
     header=F,col.names=c("wind","solrad","CO","NO","NO2","O3","HC"))
attach(prob1.6)

plot(prob1.6)

X <- as.matrix(prob1.6)
(xbar <- as.matrix(colMeans(X),ncol=1))

n <- nrow(X)

(S_n <- ((n-1)/n)*cov(X))

(R_n <- cor(X))
```

R Text Output

```
> X <- as.matrix(prob1.6)
> (xbar <- as.matrix(colMeans(X),ncol=1))
           [,1]
wind    7.500000
solrad 73.857143
CO      4.547619
NO      2.190476
NO2    10.047619
O3      9.404762
HC      3.095238
>
> n <- nrow(X)
>
> (S_n <- ((n-1)/n)*cov(X))
             wind      solrad         CO         NO        NO2         O3        HC
wind    2.4404762  -2.7142857 -0.3690476 -0.4523810 -0.5714286 -2.1785714 0.1666667
solrad -2.7142857 293.3605442  3.8163265 -1.3537415  6.6020408 30.0578231 0.6088435
CO     -0.3690476   3.8163265  1.4858277  0.6575964  2.2596372  2.7545351 0.1383220
NO     -0.4523810  -1.3537415  0.6575964  1.1541950  1.0623583 -0.7913832 0.1723356
NO2    -0.5714286   6.6020408  2.2596372  1.0623583 11.0929705  3.0521542 1.0192744
O3     -2.1785714  30.0578231  2.7545351 -0.7913832  3.0521542 30.2409297 0.5804989
HC      0.1666667   0.6088435  0.1383220  0.1723356  1.0192744  0.5804989 0.4671202
>
> (R_n <- cor(X))
             wind      solrad         CO          NO        NO2         O3         HC
wind    1.0000000 -0.10144191 -0.1938032 -0.26954261 -0.1098249 -0.2535928 0.15609793
solrad -0.1014419  1.00000000  0.1827934 -0.07356907  0.1157320  0.3191237 0.05201044
CO     -0.1938032  0.18279338  1.0000000  0.50215246  0.5565838  0.4109288 0.16603235
NO     -0.2695426 -0.07356907  0.5021525  1.00000000  0.2968981 -0.1339521 0.23470432
NO2    -0.1098249  0.11573199  0.5565838  0.29689814  1.0000000  0.1666422 0.44776780
O3     -0.2535928  0.31912373  0.4109288 -0.13395214  0.1666422  1.0000000 0.15445056
HC      0.1560979  0.05201044  0.1660323  0.23470432  0.4477678  0.1544506 1.00000000
>
```

1.5



1.9

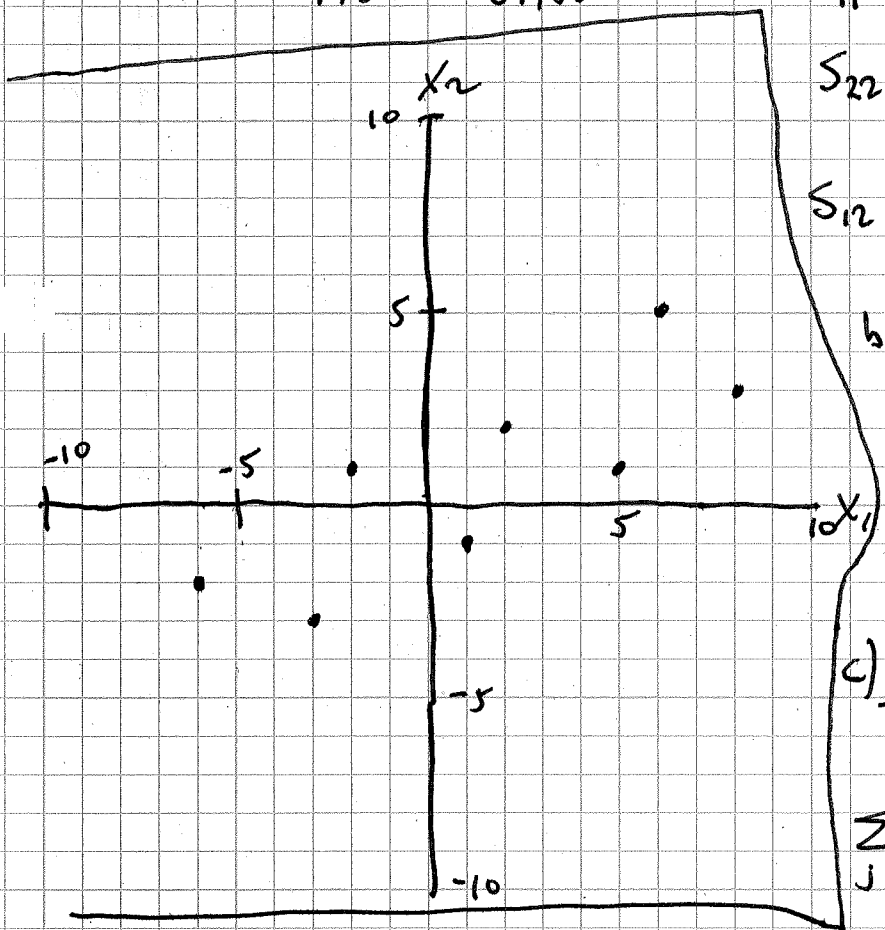
a)

x_1	x_2	x_1^2	x_2^2	x_1x_2	$\tilde{x}_1$	$\tilde{x}_2$
-6	-2	36	4	12	-6.270	0.830
-3	-3	9	9	9	-4.011	-1.383
-2	1	4	1	-2	-1.360	1.775
1	-1	1	1	-1	0.461	-1.337
2	2	4	4	4	2.674	0.922
5	1	25	1	5	4.933	-1.291
6	5	36	25	30	7.584	1.867
8	3	64	9	24	8.506	-0.807
Sum	11	6	179	54	81	12.517	0.576
Mean	1.375	0.750					

$$s_{11} = \frac{1}{8}\left[179 - \frac{11^2}{8}\right] = 20.4844$$

$$s_{22} = \frac{1}{8}\left[54 - \frac{6^2}{8}\right] = 6.1875$$

$$s_{12} = \frac{1}{8}\left[81 - \frac{11(6)}{8}\right] = 9.0938$$



b) $\tilde{x}_1 = x_1\cos(26) + x_2\sin(26)$

$= 0.899x_1 + 0.438x_2$

$\tilde{x}_2 = -x_1\sin(26) + x_2\cos(26)$

$= -0.438x_1 + 0.899x_2$

c) $\sum_j \tilde{x}_{j1}^2 = (-6.270)^2 + \cdots + (8.506) = 218.8170$

$\sum_j \tilde{x}_{j2}^2 = (0.830)^2 + \cdots + (-0.807)^2 = 14.1935$

$$\tilde{s}_{11} = \frac{1}{8}\left[218.8170 - \frac{12.517^2}{8}\right] = 24.9041$$

$$\tilde{s}_{22} = \frac{1}{8}\left[14.1935 - \frac{0.576^2}{8}\right] = 1.7690$$

1.9 d) $(x_1, x_2) = (4, -2) \Rightarrow \tilde{x}_1 = .899(4) + .438(-2) = 2.72$

$\tilde{x}_2 = -.438(4) + .899(-2) = -3.55$

$P = (\tilde{x}_1, \tilde{x}_2) = (2.72, -3.55)$

Equation 1-17: $d(O,P) = \sqrt{\frac{\tilde{x}_1^2}{s_{11}} + \frac{\tilde{x}_2^2}{s_{22}}} = \sqrt{\frac{(2.72)^2}{24.9041} + \frac{(-3.55)^2}{1.7690}}$

$= \sqrt{7.4212} = 2.7242$

e) Equation 1-19 $d(O,P) = \sqrt{a_{11}x_1^2 + a_{22}x_2^2 + 2a_{12}x_1x_2}$

$\theta = 26$: $\cos^2\theta = .899^2 = .8082$ $\sin^2\theta = .438^2 = .1918$ $\cos\theta\sin\theta = .899(.438) = .3938$

$\cos^2\theta s_{11} + 2\sin\theta\cos\theta s_{12} + \sin^2\theta s_{22} =$

$.899^2(20.4844) + 2(.899)(.438)(9.0938) + .438^2(6.1875) = 24.9041$

$\cos^2\theta s_{22} - 2\sin\theta\cos\theta s_{12} + \sin^2\theta s_{11}$

$= .899^2(6.1875) - 2(.438)(.899)(9.0938) + .438^2(20.4844) = 1.7690$

$a_{11} = \frac{.8082}{24.9041} + \frac{.1918}{1.7690} = .1409$

$a_{22} = \frac{.1918}{24.9041} + \frac{.8082}{1.7690} = .4646$

$a_{12} = \frac{.3938}{24.9041} - \frac{.3938}{1.7690} = -.2068$

$d(O,P) = \sqrt{.1409(4)^2 + .4646(-2)^2 + 2(-.2068)(4)(-2)}$

$= \sqrt{7.4216} = 2.7243$

1.11 $a_{11} = 4 \quad a_{22} = 1 \quad a_{12} = -1$

Eq. 1-20: $d(P,Q) = \sqrt{a_{11}(x_1-y_1)^2 + 2a_{12}(x_1-y_1)(x_2-y_2) + a_{22}(x_2-y_2)^2}$

Inner part: $4(x_1-y_1)^2 - 2(x_1-y_1)(x_2-y_2) + (x_2-y_2)^2$

$= 3(x_1-y_1)^2 + [(x_1-y_1)^2 - 2(x_1-y_1)(x_2-y_2) + (x_2-y_2)]^2$

$= 3(x_1-y_1)^2 + [(x_1-y_1) - (x_2-y_2)]^2$

1) $d(P,Q) = d(Q,P)$

$d(Q,P) = 3(y_1-x_1)^2 + [(y_1-x_1) - (y_2-x_2)]^2$

$= 3(x_1-y_1)^2 + [-(x_1-y_1) - (-(x_2-y_2))]^2$

$= 3(x_1-y_1)^2 + (-1)^2[(x_1-y_1) - (x_2-y_2)]^2 = d(P,Q)$

2) $d(P,Q) > 0$ if $P \neq Q$. $\Rightarrow x_1 \neq y_1$ and/or $x_2 \neq y_2$

$\Rightarrow d(Q,P) > 0 \quad \Rightarrow (x_1-y_1)^2 > 0$ and/or $(x_2-y_2)^2 > 0$

3) $d(P,Q) = 0 \Rightarrow P = Q \Rightarrow x_1 = y_1, x_2 = y_2$

Clearly $3(0)^2 + [0-0]^2 = 0$

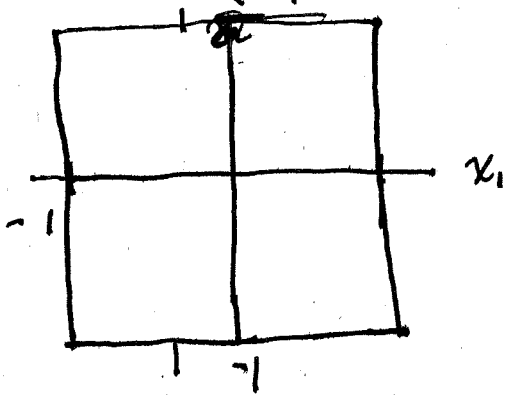
1.12 $P = (x_1, x_2)$ $O = (0, 0)$

$d(O, P) = \max(|x_1|, |x_2|)$

a) $P = (-3, 4) \Rightarrow |x_1| = 3, \; |x_2| = 4$

$\Rightarrow d(O, P) = 4$

b)



Locus of Points w/ $d(O, P) = 1$

$\{(x_1, x_2)\}$ s.t. $\max(|x_1|, |x_2|) = 1$

c) In p-dimensions hyper-cube

1.10

Problem 1.15.

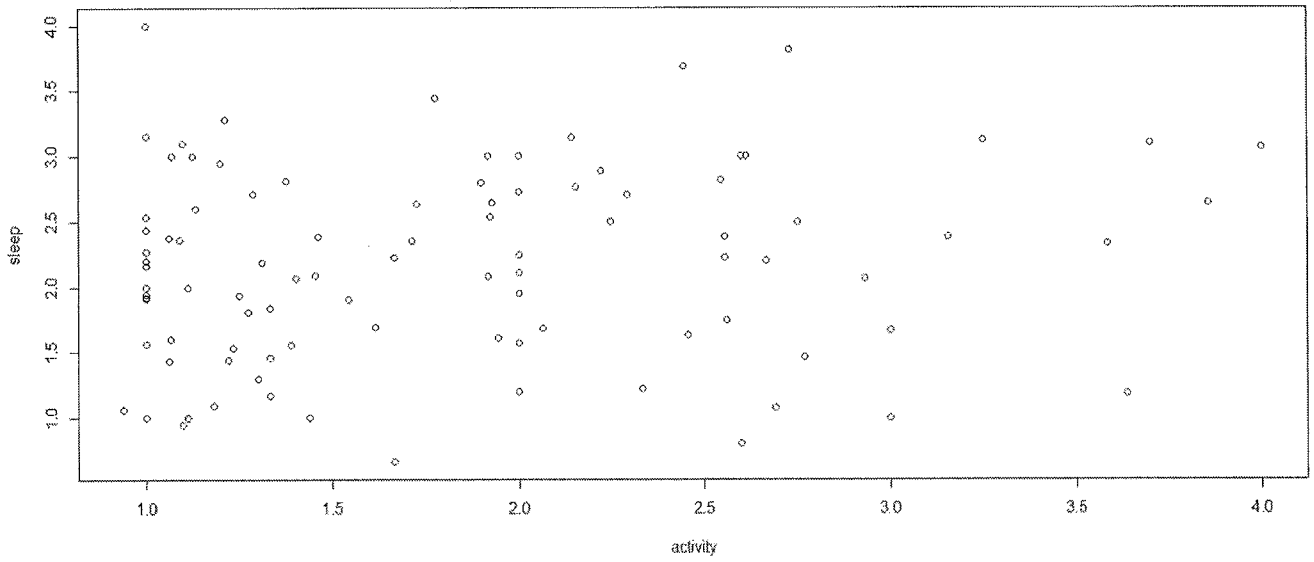
R Program

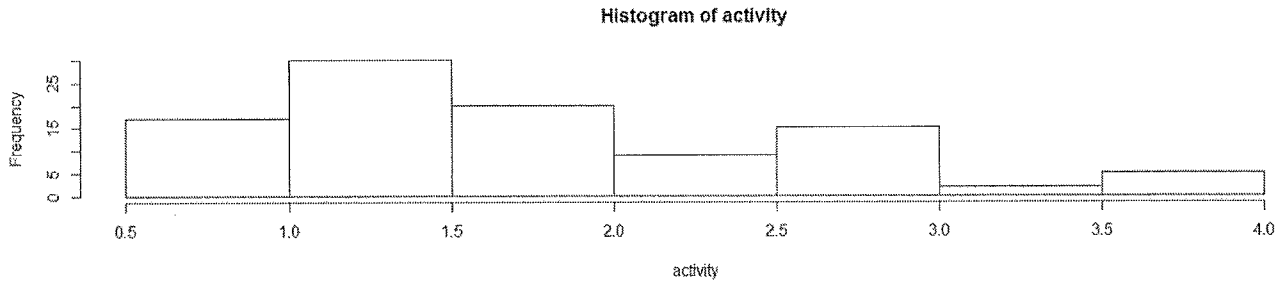
```
prob1.15 <- read.table("http://www.stat.ufl.edu/~winner/sta4702/data/wichern/T1-
7.DAT",
    header=F, col.names=c("symptoms","activity","sleep","eat","appetite","skin"))
attach(prob1.15)

plot(activity,sleep)
par(mfrow=c(2,1))
hist(activity)
hist(sleep)

n <- length(symptoms)

X <- as.matrix(prob1.15)

(xbar <- matrix(colMeans(X),ncol=1))

(S_n <- ((n-1)/n) * cov(X))
(R <- cor(X))
```

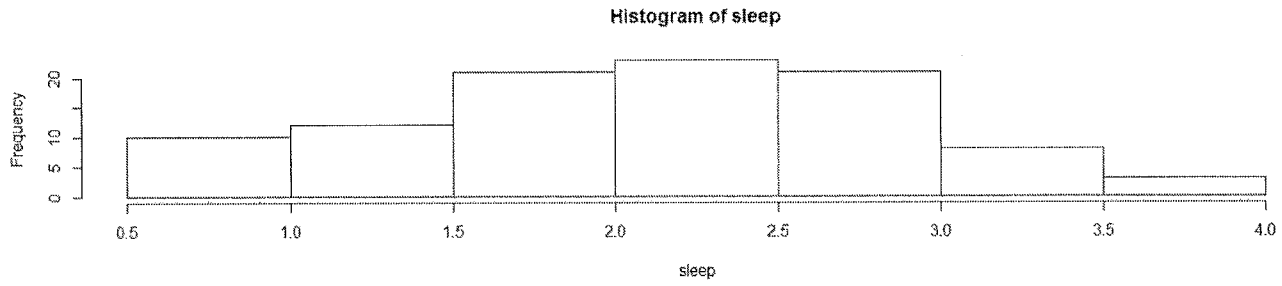
R Text Output

```
> (xbar <- matrix(colMeans(X),ncol=1))
         [,1]
[1,] 3.542347
[2,] 1.809357
[3,] 2.137602
[4,] 2.209000
[5,] 2.574827
[6,] 1.275510
>
> (S_n <- ((n-1)/n) * cov(X))
          symptoms   activity      sleep        eat     appetite         skin
symptoms 4.6072534  0.9218418 0.58368175 0.27408964  1.063917438  0.156537068
activity 0.9218418  0.6065679 0.10980144 0.11726018  0.384918205 -0.024598397
sleep    0.5836818  0.1098014 0.56559795 0.08611715  0.344438992  0.109007601
eat      0.2740896  0.1172602 0.08611715 0.10928245  0.215187224  0.021591837
appetite 1.0639174  0.3849182 0.34443899 0.21518722  0.853374694 -0.008727718
skin     0.1565371 -0.0245984 0.10900760 0.02159184 -0.008727718  0.852665556
> (R <- cor(X))
           symptoms    activity     sleep        eat    appetite         skin
symptoms 1.00000000  0.55143669 0.3615773 0.38627479  0.53655840  0.07897812
activity 0.55143669  1.00000000 0.1874625 0.45544470  0.53500626 -0.03420407
sleep    0.36157729  0.18746250 1.0000000 0.34638617  0.49577944  0.15696886
eat      0.38627479  0.45544470 0.3463862 1.00000000  0.70464665  0.07073348
appetite 0.53655840  0.53500626 0.4957794 0.70464665  1.00000000 -0.01023155
skin     0.07897812 -0.03420407 0.1569689 0.07073348 -0.01023155  1.00000000
>
```

Problem 1.15







Note that both activity and sleep appear to have values below 1.

1.12

Problem 1.17

R Program

```
prob1.17 <- read.csv("E:\\orangedrive\\sta4702\\table1_9.csv")
attach(prob1.17)

n <- length(m100)

X <- as.matrix(prob1.17[,2:8])

(xbar <- matrix(colMeans(X),ncol=1))

(S_n <- ((n-1)/n) * cov(X))
(R <- cor(X))
```

R Output

```
> (xbar <- matrix(colMeans(X),ncol=1))
           [,1]
[1,]  11.357778
[2,]  23.118519
[3,]  51.989074
[4,]   2.022407
[5,]   4.189444
[6,]   9.080741
[7,] 153.619259
>
> (S_n <- ((n-1)/n) * cov(X))
             m100       m200       m400        m800       m1500       m3000      m42195
m100    0.15243951  0.3381800  0.8747905 0.027190535 0.08233765  0.22955165   4.253915
m200    0.33818004  0.8471052  2.1522283 0.064940604 0.19900844  0.54408443  10.192673
m400    0.87479053  2.1522283  6.6205417 0.178441118 0.49974763  1.40039328  28.368477
m800    0.02719053  0.0649406  0.1784411 0.007407167 0.02101800  0.06024266   1.197068
m1500   0.08233765  0.1990084  0.4997476 0.021018004 0.07280895  0.21215226   3.474285
m3000   0.22955165  0.5440844  1.4003933 0.060242661 0.21215226  0.65244760  10.507830
m42195 4.25391502 10.1926730 28.3684771 1.197068450 3.47428477 10.50783018 265.265148
> (R <- cor(X))
             m100      m200      m400      m800      m1500      m3000     m42195
m100    1.0000000 0.9410886 0.8707802 0.8091758 0.7815510 0.7278784 0.6689597
m200    0.9410886 1.0000000 0.9088096 0.8198258 0.8013282 0.7318546 0.6799537
m400    0.8707802 0.9088096 1.0000000 0.8057904 0.7197996 0.6737991 0.6769384
m800    0.8091758 0.8198258 0.8057904 1.0000000 0.9050509 0.8665732 0.8539900
m1500   0.7815510 0.8013282 0.7197996 0.9050509 1.0000000 0.9733801 0.7905565
m3000   0.7278784 0.7318546 0.6737991 0.8665732 0.9733801 1.0000000 0.7987302
m42195 0.6689597 0.6799537 0.6769384 0.8539900 0.7905565 0.7987302 1.0000000
```

1.13

Problem 1.18

R Program

```
prob1.18 <- read.csv("E:\\orangedrive\\sta4702\\table1_9.csv")
attach(prob1.18); names(prob1.18)

s100 <- 100/m100; s200 <- 200/m200; s400 <- 400/m400
s800 <- 800/(m800*60); s1500 <- 1500/(m1500*60); s3000 <- 3000/(m3000*60)
s42195 <- 42195/(m42195*60)

n <- length(s100)

X <- cbind(s100,s200,s400,s800,s1500,s3000,s42195)

(xbar <- matrix(colMeans(X),ncol=1))

(S_n <- ((n-1)/n) * cov(X))
(R <- cor(X))
```

R Output

```
> (xbar <- matrix(colMeans(X),ncol=1))
         [,1]
[1,] 8.814772
[2,] 8.664408
[3,] 7.712067
[4,] 6.604214
[5,] 5.989687
[6,] 5.542701
[7,] 4.620264
>
> (S_n <- ((n-1)/n) * cov(X))
             s100       s200       s400       s800       s1500      s3000      s42195
s100   0.08886163 0.09383586 0.09488221 0.06385913 0.08069721 0.09043587 0.07959802
s200   0.09383586 0.11254789 0.11176120 0.07353737 0.09424082 0.10348392 0.09158236
s400   0.09488221 0.11176120 0.13523721 0.07944200 0.09367553 0.10631059 0.09999405
s800   0.06385913 0.07353737 0.07944200 0.07216131 0.08485322 0.09790735 0.09255923
s1500  0.08069721 0.09424082 0.09367553 0.08485322 0.12154716 0.14105343 0.11626411
s3000  0.09043587 0.10348392 0.10631059 0.09790735 0.14105343 0.17331425 0.14384635
s42195 0.07959802 0.09158236 0.09999405 0.09255923 0.11626411 0.14384635 0.16362679
> (R <- cor(X))
            s100      s200      s400      s800      s1500     s3000     s42195
s100   1.0000000 0.9383028 0.8655248 0.7974687 0.7764777 0.7287297 0.6601124
s200   0.9383028 1.0000000 0.9058875 0.8159945 0.8057456 0.7409469 0.6748635
s400   0.8655248 0.9058875 1.0000000 0.8041737 0.7306437 0.6944025 0.6722005
s800   0.7974687 0.8159945 0.8041737 1.0000000 0.9060324 0.8754795 0.8518052
s1500  0.7764777 0.8057456 0.7306437 0.9060324 1.0000000 0.9718385 0.8244153
s3000  0.7287297 0.7409469 0.6944025 0.8754795 0.9718385 1.0000000 0.8541900
s42195 0.6601124 0.6748635 0.6722005 0.8518052 0.8244153 0.8541900 1.0000000
```