C-A-2

ASANSOL ENGINEERING COLLEGE

NAME: PRAKRITI MUKHERJEE

DEPT: CSBS

Roll NO: 10831122008

SUB NAME: COMPUTATIONAL STATISTICS

SUB CODE: BSC-391

(1) a) The scatter plot of the data & marginal dot diagram are illustrated below

b) The coordinates of the points are determined by paired measurements:
(1, 18.95), (2, 19.00), (3, 17.95), (3, 15.54), (4, 14.00), (5, 12.95), (6, 8.94), (8, 7.49), (9, 6.00).
As we can see small values of x_1 occurs with large values of x_2 & vice versa. So, the sign of sample covariance s_{12} will be negative

c) $\bar{x}_1 = \frac{1}{10}\sum_{j=1}^{10} x_{j1} = \frac{1}{10}(1+2+3+3+4+5+6+8+9+11) = 5.2$

$\bar{x}_2 = \frac{1}{10}\sum_{j=1}^{10} x_{j2} = \frac{1}{10}(18.95+19.00+17.95+15.54+14.00+12.95+8.94+7.49+6.00+3.99) = 12.481$

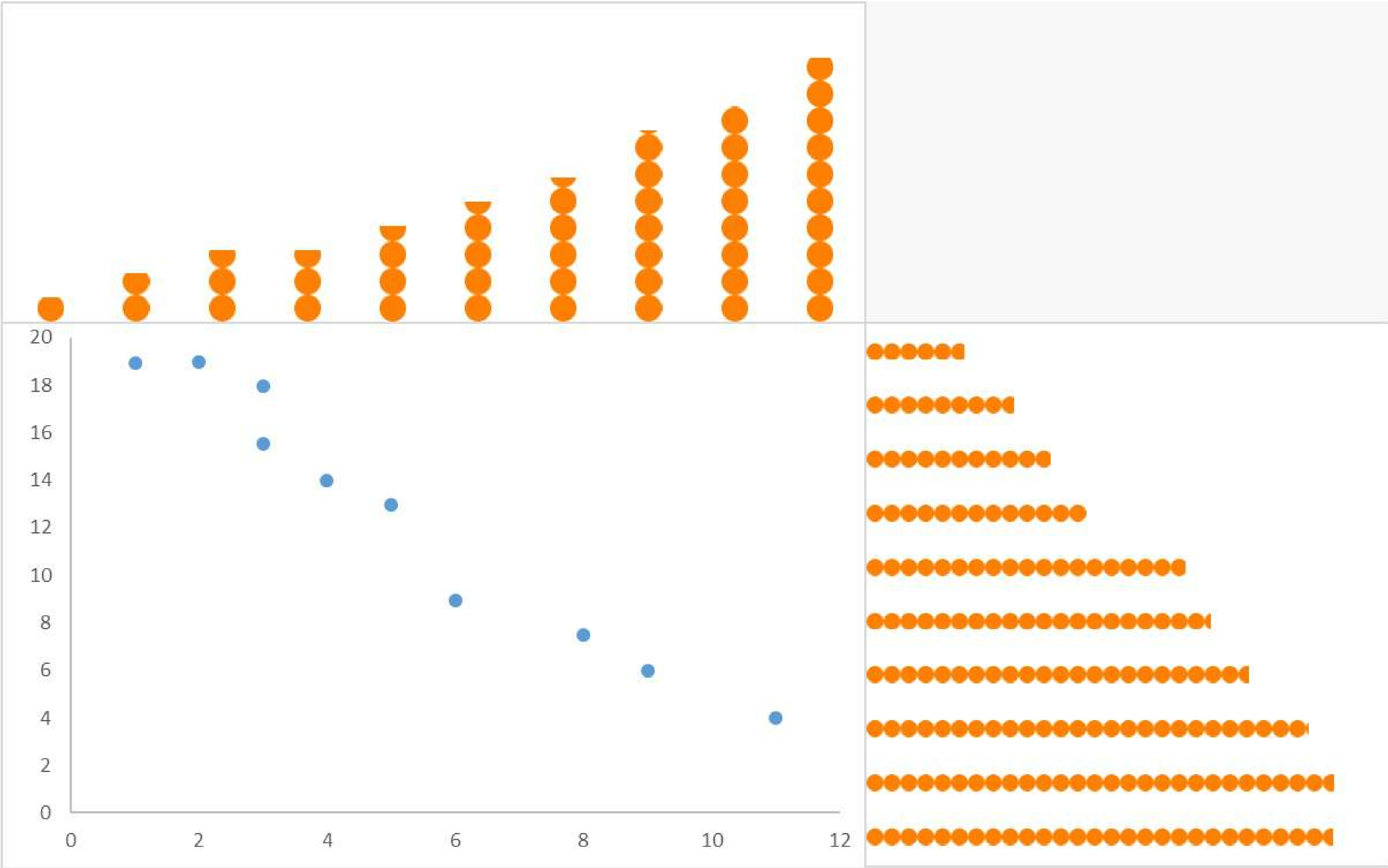
$s_{11} = \frac{1}{10}\sum_{j=1}^{10}(x_{j1}-\bar{x}_1)^2 = \frac{1}{10}\Big[(1-5.2)^2+(2-5.2)^2+(3-5.2)^2+(3-5.2)^2+(4-5.2)^2$
$(5-5.2)^2+(6-5.2)^2+(8-5.2)^2+(9-5.2)^2+(11-5.2)^2\Big] = 9.56$

$s_{22} = \frac{1}{10}\sum_{j=1}^{10}(x_{j2}-\bar{x}_2)^2 = \frac{1}{10}\Big[(18.95-12.481)^2+(19.00-12.481)^2+$
$(17.95-12.481)^2+(15.54-12.481)^2+(14.00-12.481)^2+$
$(12.95-12.481)^2+(8.94-12.481)^2+(7.49-12.481)^2+(6.00-12.481)^2+$
$(3.99-12.481)^2\Big] = 27.769$

$s_{12} = \frac{1}{10}\sum_{j=1}^{10}(x_{j1}-\bar{x}_1)(x_{j2}-\bar{x}_2) = \frac{1}{10}\Big[(1-5.2)(18.95-12.481)+$
$(2-5.2)(19.00-12.481)+(3-5.2)(17.95-12.481)+(3-5.2)$
$(15.54-12.481)+(4-5.2)(14.00-12.481)+(5-5.2)(12.95-12.481)$
$+(6-5.2)(8.94-12.481)+(8-5.2)(7.49-12.481)+(9-5.2)(6-12.481)$
$+(11-5.2)(3.99-12.481) = -15.939$

$r_{12} = \frac{s_{12}}{\sqrt{s_{11}}\sqrt{s_{22}}} = \frac{-15.939}{\sqrt{9.56}\sqrt{27.769}} = -0.978$

The negative sign of r_{12} indicates that x_1 is increasing, x_2 is decreasing & vice versa. Also r_{12} is close to -1, so correlation is strong


20
18
16
14
12
10
8
6
4
2
0
0
2
4
6
8
10
12

d) since $s_{12} = s_{21}$, $r_{12} = r_{21}$, we get

$$\bar{X} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} 5.2 \\ 12.481 \end{bmatrix}$$

$$S_n = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} = \begin{bmatrix} 9.56 & -15.939 \\ -15.939 & 27.769 \end{bmatrix}$$

$$R = \begin{bmatrix} 1 & r_{12} \\ r_{21} & 1 \end{bmatrix} = \begin{bmatrix} 1 & -0.978 \\ -0.978 & 1 \end{bmatrix}$$

2 a) The scatter diagram & marginal dot for variables x_1 & x_2 are illustrated:

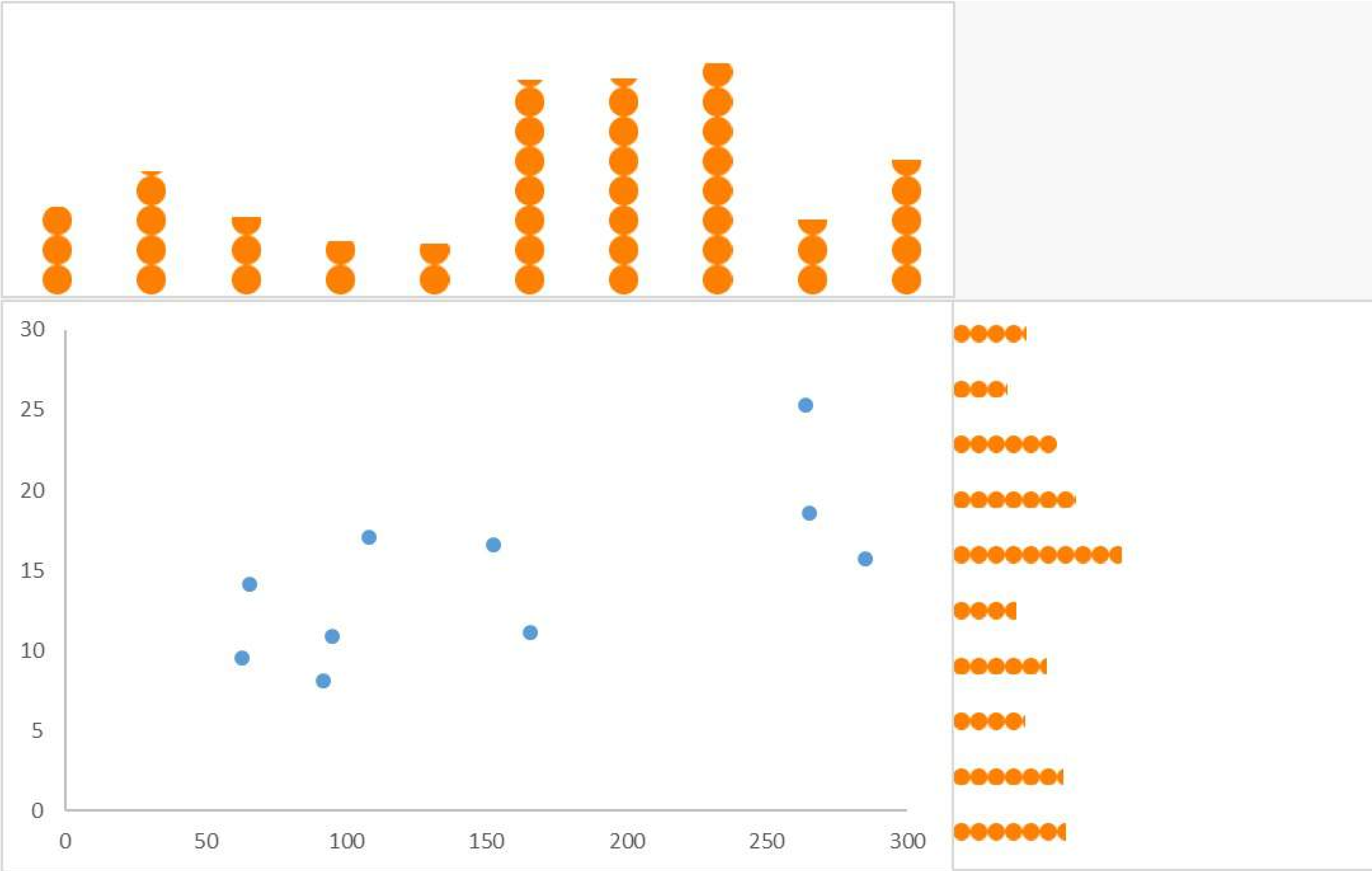
As we can see the blue points appear randomly scattered on the coordinate plane. That means that there is a weak correlation between the variables (if large correlation exists, the points concentrate near a straight line)

b) $$\bar{x}_1 = \frac{1}{10}\sum_{j=1}^{10} x_{j1} = \frac{1}{10}(108.28 + 152.36 + 95.04 + 65.45 + 62.97 + 263.99 + 265.19 + 285.06 + 92.01 + 165.68)$$
$$= 155.603$$

$$\bar{x}_2 = \frac{1}{10}\sum_{j=1}^{10} x_{j2} = \frac{1}{10}[(17.05 + 16.59 + 10.91 + 14.14 + 9.52 + 25.33 + 18.54 + 15.73 + 8.10 + 11.13)] = 14.704$$

$$s_{11} = \frac{1}{10}\sum_{j=1}^{10}(x_{j1} - \bar{x}_1)^2 = \frac{1}{10}[(108.28 - 155.603)^2 + (152.36 - 155.603)^2 + (95.04 - 155.603)^2 + (65.45 - 155.603)^2 + (62.97 - 155.603)^2 + (263.99 - 155.603)^2 + (265.19 - 155.603)^2 + (285.06 - 155.603)^2 \; (92.01 - 155.603)^2 + (165.68 - 155.603)^2 = 6728.808$$

$$s_{22} = \frac{1}{10}\sum_{j=1}^{10}(x_{j2} - \bar{x}_2)^2 = \frac{1}{10}[(17.05 - 14.704)^2 + (16.59 - 14.704)^2 + (10.91 - 14.704)^2 + (14.14 - 14.704)^2 + (9.52 - 14.704)^2 + (25.33 - 14.704)^2 + (18.54 - 14.704)^2 + (15.73 - 14.704)^2 + (8.10 - 14.704)^2 + (11.13 - 14.704)^2 = 23.571$$



30
25
20
15
10
5
0
0
50
100
150
200
250
300

$$s_{12} = \frac{1}{10}\sum_{j=1}^{10}(x_{j1}-\bar{x}_1)(x_{j2}-\bar{x}_2) = \frac{1}{10}\Big[(108.28-155.603)(17.05-14.704) +$$
$$(152.36-155.603)(16.59-14.704) + (95.04-155.603)(10.91-14.704)$$
$$+(65.45-155.603)(14.14-14.704) + (62.97-155.603)(9.52-14.70) +$$
$$(263.99-155.603)(25.33-14.704) + (265.19-155.603)(18.54-14.704) +$$
$$(285.06-155.603)(15.73-14704) + (92.01-155.603)(8.10-14.704) +$$
$$(165.68-155.603)(11.13-14.704) = 273.257$$

$$r_{12} = \frac{s_{12}}{\sqrt{s_{11}}\sqrt{s_{22}}} = \frac{273.257}{\sqrt{6728.808}\sqrt{23.571}} = 0.686$$

The positive sign of r_{12} indicates that as the value of variable x_1 (sales) increases, the value of variable x_2 (profits) increases. Also, as values of one variable decreases, the value of the other one decreases. So $r_{12} = 0.686$, the correlation is weak.

(2) a> The scatter diagram & marginal dot for (x_2, x_3) & (x_1, x_3) are illustrated:

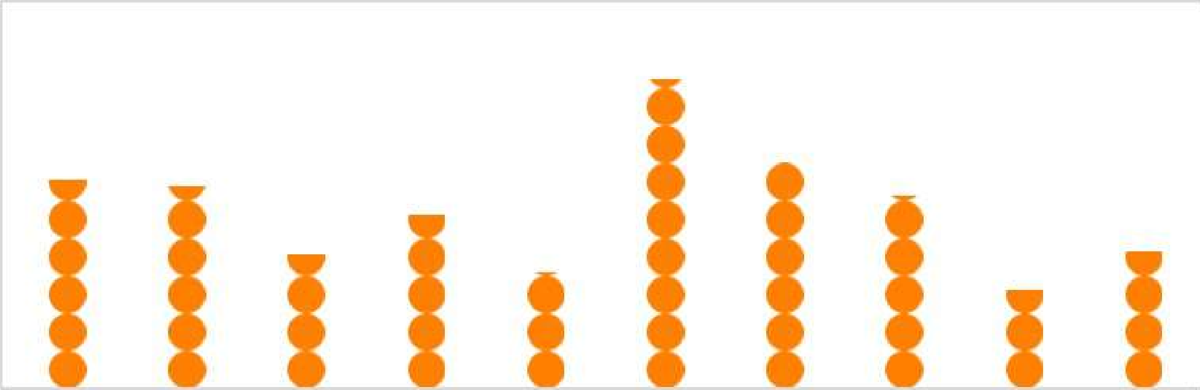
There is negative correlation between x_2 & x_3 & negative correlation between x_1 & x_3. The marginal distribution of x_1 appears to be skewed to the right. The marginal distribution of x_2 seems reasonably symmetric. The marginal distribution of x_3 also appears to be skewed to the right.

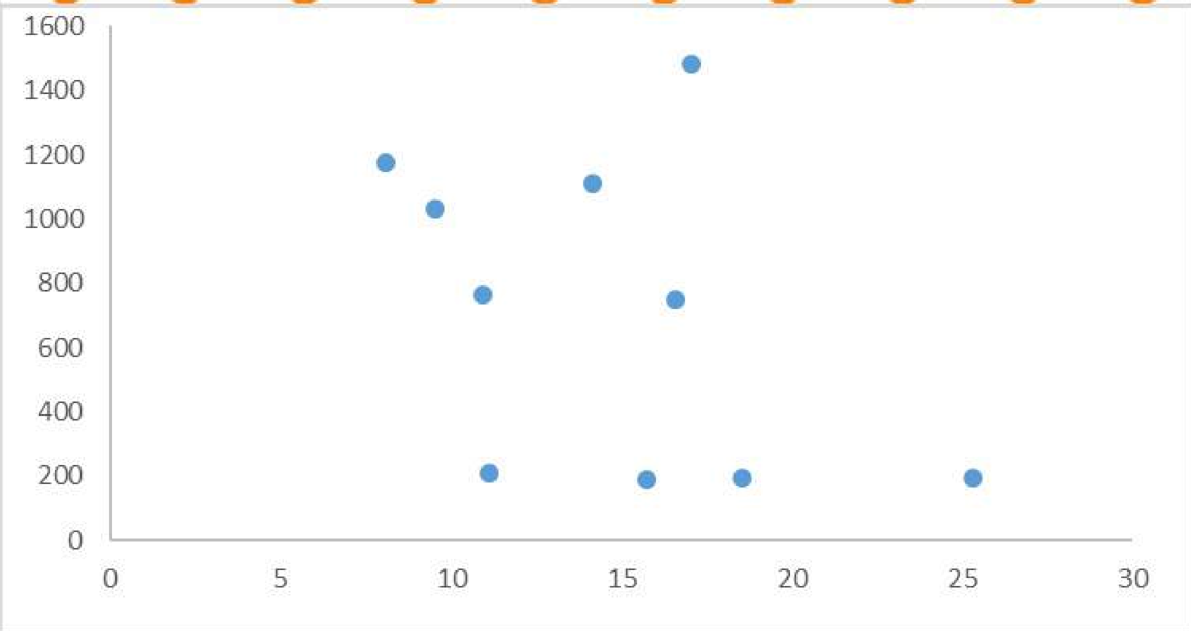
b> $$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} = \begin{bmatrix} 155.603 \\ 14.704 \\ 710.911 \end{bmatrix}$$

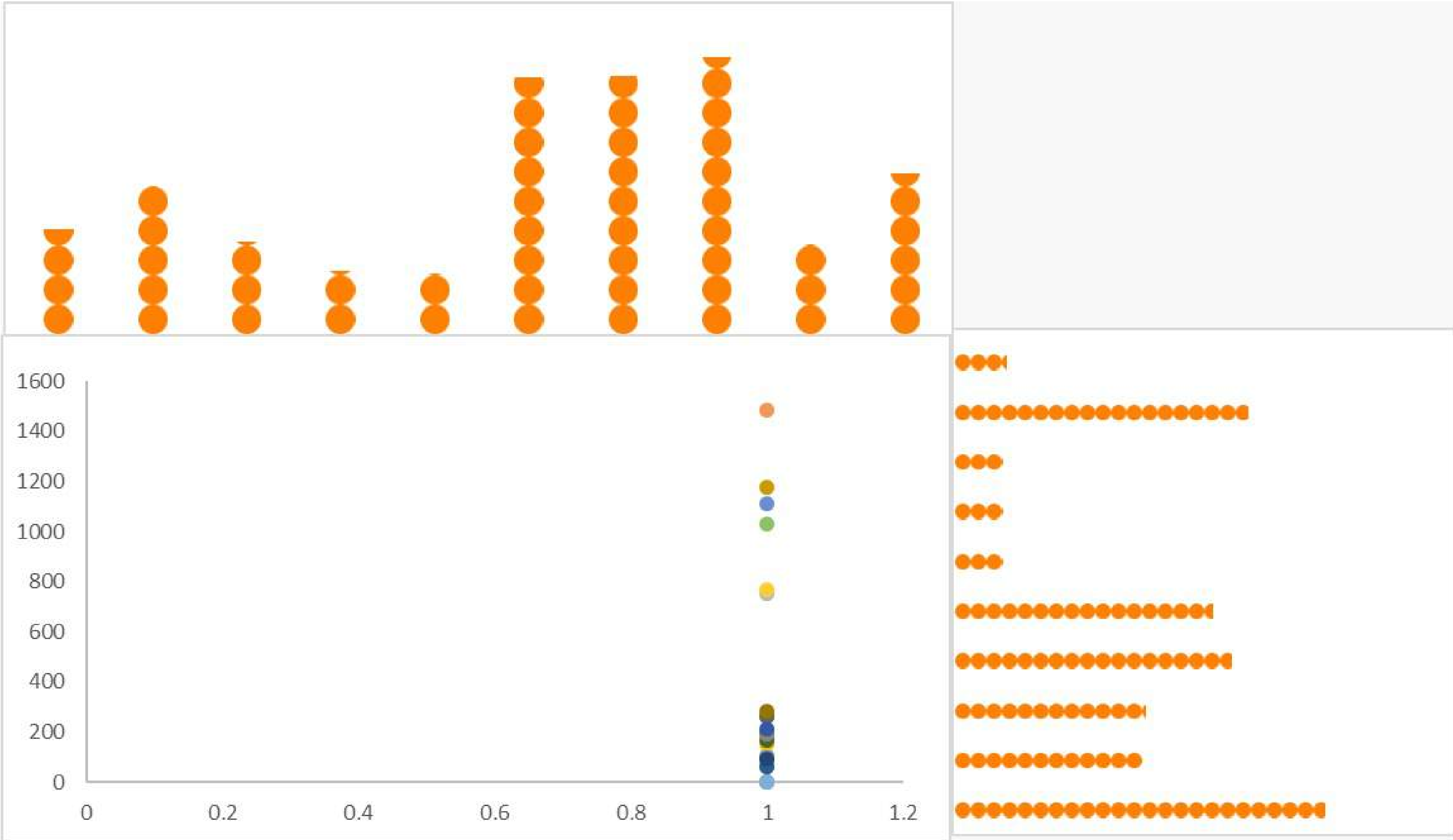
$$S_n = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix} = \begin{pmatrix} 6728.808 & 273.257 \\ 273.257 & 23.57 \end{pmatrix}$$

$$R = \begin{bmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{bmatrix}$$

$$R = \begin{bmatrix} 1 & 0.686 & -0.845 \\ 0.686 & 1 & -0.845 \\ -0.845 & -0.845 & 1 \end{bmatrix}$$




1600
1400
1200
1000
800
600
400
200
0
0
5
10
15
20
25
30


1600
1400
1200
1000
800
600
400
200
0
0
0.2
0.4
0.6
0.8
1
1.2