Assignment - 1

Answer-1, PDF of Normal distribution -

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Here, $\mu = 42$ months, $\sigma = 8$ months

we need probability for $20 \le x \le 30$.

$$P(20 \le x \le 30) = \int_{20}^{30} f(x)\,dx$$

$$\int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx = \phi\left(\frac{b-\mu}{\sigma}\right) - \phi\left(\frac{a-\mu}{\sigma}\right)$$

where,

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\,dt$$

here, $z = \frac{X-\mu}{\sigma}$

$$z_1 = \frac{20-42}{8} = \frac{-22}{8} = -2.75$$

$$z_2 = \frac{30-42}{8} = \frac{-12}{8} = -1.5$$

$$P(20 \le x \le 30) = P(-2.75 \le z \le -1.5)$$

$$= \phi(-1.5) - \phi(-2.75)$$

$$= 0.0668 - 0.0030$$

$$\boxed{P = 0.0638}$$

Date :
Page :

Answer-2 Yes, trimmed mean is a meaningful measure of location for this data because we don't know what is the exact value of 100* and also 36 is very less value we can take them as outliers.

$$\text{trimmed mean (TM)} = \frac{\sum_{n=p+1}^{N-p} x_n}{N-2p}$$

Here $p = 1$

arrange data in increasing order →

36 [45, 51, 63, 75, 80, 90,] 100*

$$TM = \frac{45+51+63+75+80+90}{6}$$

$$= \frac{404}{6}$$

$$\boxed{TM = 67.33 \text{ hr}}$$

Date :
Page :

Answer-3 →

Arrange data in increasing order -

1.75, 1.91, 1.92, 2.35, 2.53, 2.68, 3.09, 3.15

- mean = $\frac{1.75+1.91+1.92+2.35+2.53+2.62+3.09+3.15}{8}$

Sample mean = 2.415

Variance = $\frac{(1.75-2.415)^2+(1.91-2.415)^2+(1.92-2.415)^2+\cdots\cdots(3.09-2.415)+(2.415-3.15)^2}{7}$

- Sample variance = 0.2854 $= s^2$
- Standard deviation = $\sqrt{\text{Variance}}$

$= \sqrt{0.2854}$

$S = 0.5342$.

Another data →

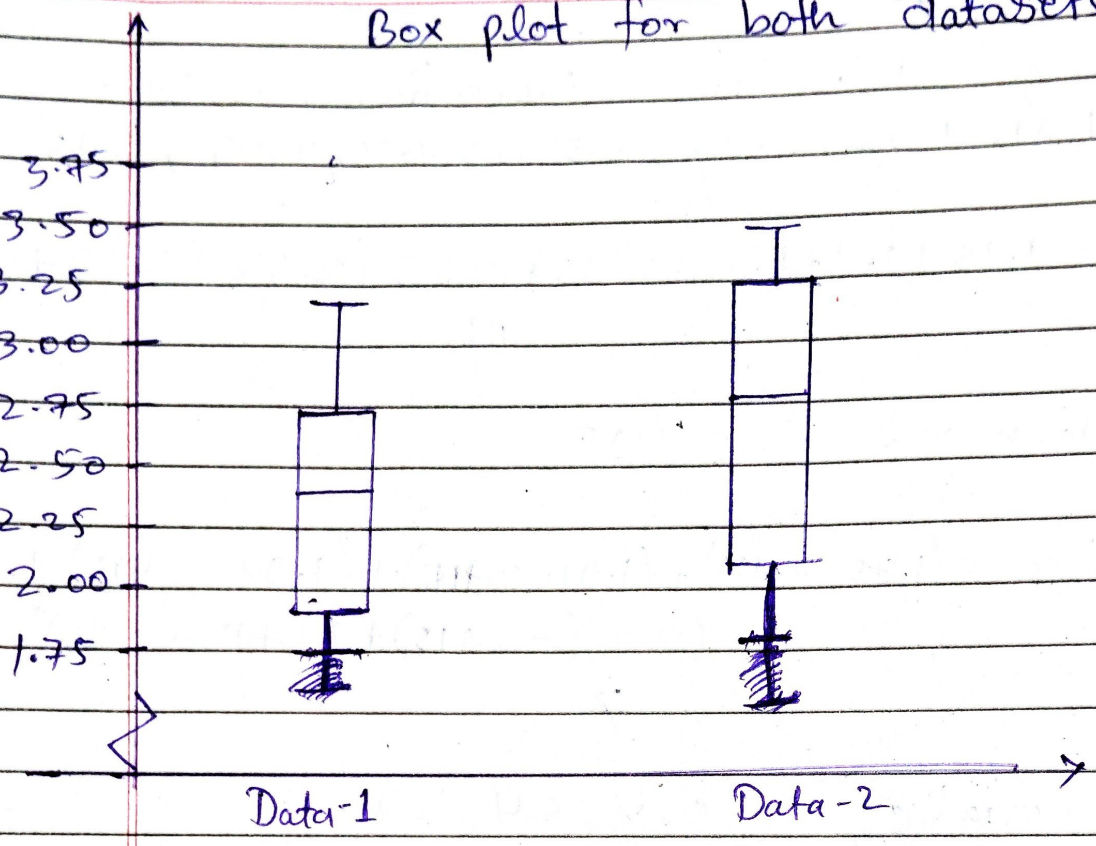
1.83, 1.89, 1.99, 2.46, 2.65, 2.87, 3.13, 3.29, 3.35, 3.4

mean = $\frac{1}{N}\sum_{n=1}^{N} x_n$ = 2.687 = $\bar{x}$

Variance = $\sigma^2 = \frac{1}{N-1}\Sigma(x-\bar{x})^2$ = 0.38329

standard deviation = $\sigma = \sqrt{0.38329} = 0.6191$

Box plot for both datasets



Q1 (25th percentile)	1.918	2.108
Q2 (50th percentile)	2.44	2.76
Q3 (75th percentile)	2.738	3.25
IQR = $(Q_3 - Q_1)$	0.82	1.142
Lower whisker $(Q_3 + \omega \times IQR)$	3.968 < let's take 3.18	4.963 < 3.48
Upper whisker $(Q_1 - \omega \times IQR)$	> 0.688 let's take 1.75	> 0.395 1.76

taking $\omega = 1.5$

Answer - 4 :-

4(a).

Weight (kg)	Normalized Wt.
50	(50-41)/95 = 0.0947
115	(115-41)/95 = 0.7789
96	(96-41)/95 = 0.5789
41	(41-41)/95 = 0
79	(79-41)/95 = 0.4
109	(109-41)/95 = 0.7157
73	(73-41)/95 = 0.3368
104	(104-41)/95 = 0.6631
64	(64-41)/95 = 0.2421
136	(136-41)/95 = 1

min-max normalization.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

here $X_{max} = 136$, $X_{min} = 41$

$X_{max} - X_{min} = 136 - 41 = 95$

Answer - 4(b).

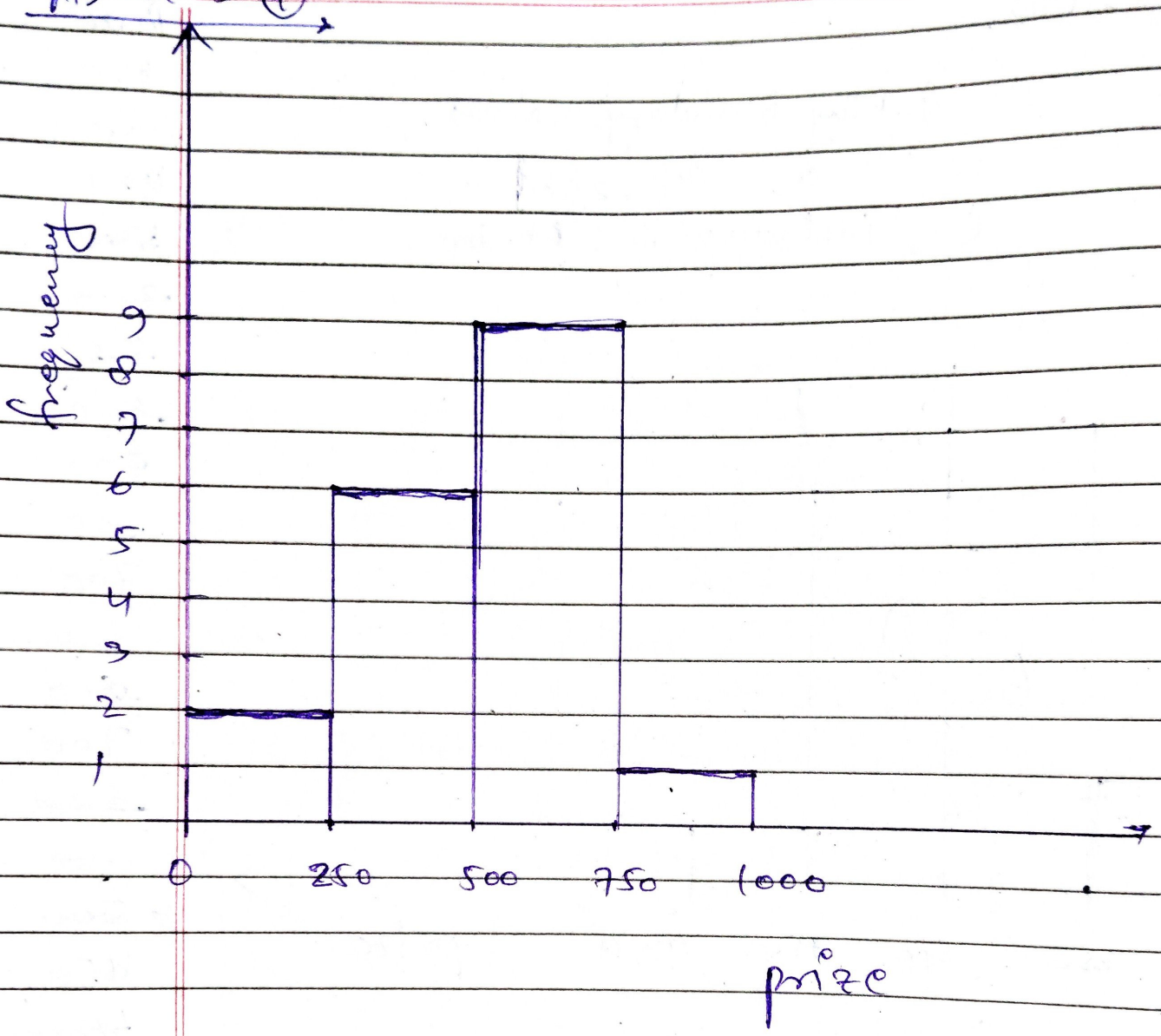
Weight (kg)	Categories
50	low
115	high
96	~~medi~~ medium
41	low
79	medium
109	high
73	medium
104	high
64	medium
136	high

Answer 4(c).

Weight(kg)	Height (m)	BMI
50	1.52	21.64
115	1.77	36.70
96	1.83	28.66
41	1.55	17.06
79	1.82	23.84
109	1.89	30.51
73	1.76	23.56
104	1.71	35.56
64	1.74	21.13
136	1.78	42.92

$$BMI = \frac{Weight}{(Height)^2}$$

Answer-5 (a)



Answer-5(b)

Store \ product	Laptop	Printer	Scanner	Desktop
New York, NY	1	2	4	3
Washington, DC	2	2	2	2

Date : Page :

Answer-5(C)

(i)

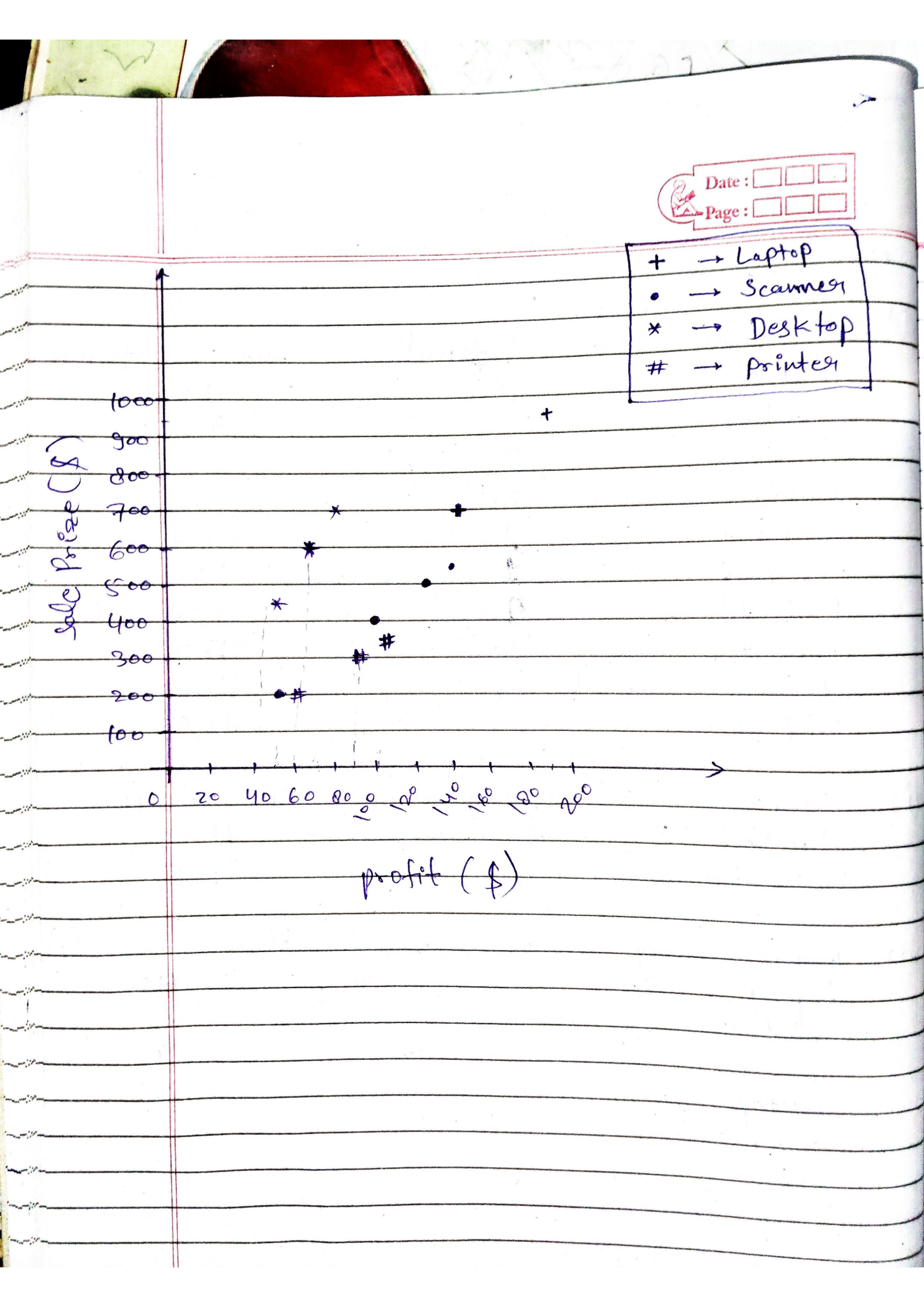
Customer	No. of. Observations	Sum of Sale Price
B. March	3	1700
J. Bain	1	500
T. Goss	2	750
L. Nye	2	900
S. Cann	1	600
E. Sims	1	700
P. Judd	2	900
G. Hinton	4	2150
M. Fu	1	450
M. Taylor	1	400

(ii)

Store	observations	Mean sale price
New York, NY	10	485
Washington, DC	8	525

(iii)

Product Catagory	observations	Sum of Profit ($)
Laptop	3	410
Printer	4	360
Scanner	6	640
Desktop	5	295


Date :
Page :
+ → Laptop
• → Scanner
* → Desktop
# → printer
Sale Price ($)
1000
900
800
700
600
500
400
300
200
100
0
20
40
60
80
100
120
140
160
180
200
profit ($)