Soft Computing

ASSIGNMENT 2

Name - PRABHAKAR KUMAR
Roll - IRM2017008

Q3

Let

$X = [2004, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017]$

$Y = [61.2, 58.3, 67.1, 69.2, 68.9, 83.5, 89.1, 80, 92.3, 93, 97]$

Let we define elements of X, as

$X = [(x - 2003) \text{ such } x \in X]$

This is just done to scale down the values of X, as them to having a larger value is insignificant

X	Y	$(X-\bar{X})(Y-\bar{Y})$	$(X-\bar{X})^2$
1	61.2	130.95	59.71
5	58.3	73.97	13.89
6	67.1	30.12	7.43
7	69.2	15.45	2.98
8	68.9	6.72	0.52
9	83.5	1.46	0.07
10	89.1	13.94	1.61
11	80	4.21	5.16
12	92.3	46.32	10.71
13	93	63.46	18.25
14	97	99.41	27.8

Here

$$\bar{x} = \frac{\sum x}{n}$$

$$= 8.72$$

$$\bar{y} = \frac{\sum y}{n}$$

$$= \underline{78.14}$$

$$\therefore m = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sum (x-\bar{x})^2}$$

$$= \frac{486.03}{148.18}$$

$$\approx \underline{3.28}$$

$$\therefore c = \bar{y} - (m * \bar{x})$$

$$= 78.14 - 8.72 \times 3.28$$

$$= \underline{49.51}$$

(a)

$$\therefore \boxed{Y = 3.28 * X + 49.51}$$

(b)

$\therefore$ For Year 2019

$$\text{Revenue} = 3.28(2019 - 2003) + 49.51$$

$$= \underline{\underline{69.20}}$$

(c)

Y	Y_Predicted	$(Y - Y\_Predicted)^2$
61.2	52.8	70.56
58.3	65.92	58.06
67.1	69.20	4.41
69.2	72.48	10.75
68.9	75.76	47.05
83.5	79.03	19.8
89.1	82.32	45.97
80	85.60	31.36
92.3	88.9	11.67
93	92.16	0.70
97	95.44	2.43

$$\therefore RMSE = \left(\frac{\sum (Y - Y\_Predicted)^2}{n} \right)^{1/2}$$

$$= \underline{5.24}$$

Q4 Given,

$X = [75, 80, 93, 65, 87, 71, 98, 68, 84, 77]$

$Y = [82, 78, 86, 72, 91, 80, 95, 72, 80, 74]$

Mean $X = \bar{X} = 79.8$

Mean $Y = \bar{Y} = 81.0$

a

X	Y	$(X-\bar{X})(Y-\bar{Y})$	$(X-\bar{X})^2$
75	82	-4.8	23.04
80	78	-0.6	0.04
93	86	66	174.24
65	72	132.2	219.
87	91	72.0	57.9
71	80	8.8	77.43
98	95	254.8	331.24
68	72	106.2	139.2
84	80	-4.2	17.64
77	74	19.6	7.83

$$\therefore m_1 = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{\Sigma(X-\bar{X})^2}$$

$$= 0.625$$

$$\therefore c_1 = \bar{Y} - (m_1 \times \bar{X})$$

$$= \underline{31.125}$$

$$\therefore \boxed{Y = 0.625 * X + 31.125}$$

(b)

Y	X	$(Y-\bar{Y})(X-\bar{X})$	$(Y-\bar{Y})^2$
82	75	-4.8	1
78	80	-0.6	9
86	93	66	25
72	65	132.2	81
91	87	72.0	100
80	71	8.8	1
95	98	254.8	196
72	68	106.2	81
80	84	-4.2	1
74	77	19.6	49

$$\therefore m_2 = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{\Sigma(Y-\bar{Y})^2}$$

$$= 1.19$$

$$C_2 = -17.13$$

$$\therefore \boxed{X = 1.19 * Y - 17.13}$$

(c) Given, ML Marks $= X = 96$

$\therefore Y = 0.625 * X + 31.125$

$= 0.625 \times 96 + 31.125$

$= 91.125$

(d) Given, HUR Marks $= Y = 95$

$\therefore X = 1.19 * Y - 17.13$

$= 1.19 \times 95 - 17.13$

$= 96.55$

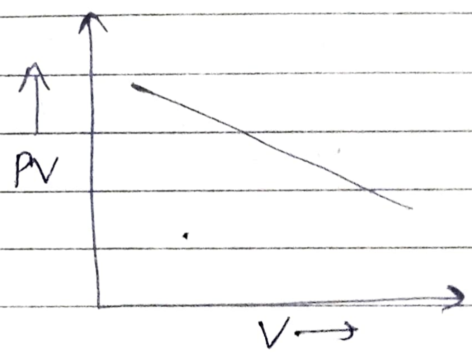
(e) We plotted the diagrams using code and scatter plots and found out the error in case of ML being independent variable was less than when we used HUR as independent variable.

Q5) Given,

V = [54.3, 61.8, 72.4, 88.7, 118.6, 194] = V

P = [61.2, 49.5, 37.5, 28.4, 19.2, 10.1] = P.

V	P	PV
54.3	61.2	3323.16
61.8	49.5	3059
72.4	37.5	2715
88.7	28.4	2519
11.8.6	19.2	2277
194	10.1	1959



Here it should be seen that as per ideal gas PV ≠ Constant, rather as V increases value of PV decreases, which shows non ideal behaviour of the gas.

This may be because the data given is recorded at different temperatures and since we do not have the value of temperature we cant find the mass of gas or the value of 'n', as

$$PV = nRT \neq \text{constant}$$

The best we can do is to fit a regression line with independent var. V and dependent var. as PV.

$\therefore$

V	PV	$(PV - \overline{PV})(\cancel{V} - \cancel{\overline{V}})$	$(PV - \overline{PV})$
54.3	3323	-29964	1936
61.8	3059	-15218	1332
72.4	2715	-1886	670
88.7	2579	1181	92
118.6	2277	-7409	412
194	1959	-65338	9158

$$\therefore m = \frac{\sum(PV - \overline{PV})(V - \overline{V})}{\sum(PV - \overline{PV})}$$

$$= -8.72$$

$c = 3499$

Here it should be noted that

$\overline{PV} = 2642$

$\overline{V} = 98.3$

$$\therefore PV = -8.72 * V + 3499$$

Hence at $V = 100$

$$P\cancel{V} = \frac{\text{Predict PV}}{100} = \underline{\underline{26.27}}$$

Q6

Given,

$X = [0, 1, 2, 3, 4, 5, 6]$

$Y = [2.4, 2.1, 3.2, 5.6, 9.3, 14.6, 21.9]$

$\therefore \sum X^0 = 7$

$\sum X^1 = 21$

$\sum X^2 = 91$

$\sum X^3 = 441$

$\sum X^4 = 2275.$

Now

$\sum (y * x^0) = 59.1$

$\sum (y * x^1) = 266.9$

$\sum (y - x^2) = 1367.5$

Now

$$\begin{bmatrix} 7 & 21 & 91 \\ 21 & 91 & 441 \\ 91 & 441 & 2275 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 59.1 \\ 266.9 \\ 1367.5 \end{bmatrix}$$

On solving the above equations using linear algebra, we get :-

$a = 2.509$

$b = -1.2$

$c = 0.7333$

$$\therefore \boxed{Y = 2.509 - 1.2 * X + 0.7333 * X^2}$$

If we find an error parameter as

$E = \sum (y - y\text{-predicted})$

$= \underline{\underline{0.163}}$

Which shows that the parabola fits the data very well.